MAY 26 2023

Distributional learning of musical pitch despite tone deafness in individuals with congenital amusia

Jiaqiang Zhu; Xiaoxiang Chen; Fei Chen; Caicai Zhang ; Jing Shao; Seth Wiener

J. Acoust. Soc. Am. 153, 3117 (2023)
https://doi.org/10.1121/10.0019472

JASA ARTICLE

Distributional learning of musical pitch despite tone deafness in individuals with congenital amusia

Jiaqiang Zhu,[1] Xiaoxiang Chen,[2,a)] Fei Chen,[2] Caicai Zhang,[1] Jing Shao,[3] and Seth Wiener[4]
[1]*Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China*
[2]*School of Foreign Languages, Hunan University, Changsha, China*
[3]*Department of English Language and Literature, Hong Kong Baptist University, Hong Kong Special Administrative Region, China*
[4]*Department of Modern Languages, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*

ABSTRACT:
Congenital amusia is an innate and lifelong deficit of music processing. This study investigated whether adult listeners with amusia were still able to learn pitch-related musical chords based on stimulus frequency of statistical distribution, i.e., via distributional learning. Following a pretest-training-posttest design, 18 amusics and 19 typical, musically intact listeners were assigned to bimodal and unimodal conditions that differed in distribution of the stimuli. Participants' task was to discriminate chord minimal pairs, which were transposed to a novel microtonal scale. Accuracy rates for each test session were collected and compared between the two groups using generalized mixed-effects models. Results showed that amusics were less accurate than typical listeners at all comparisons, thus corroborating previous findings. Importantly, amusics—like typical listeners—demonstrated perceptual gains from pretest to posttest in the bimodal condition (but not the unimodal condition). The findings reveal that amusics' distributional learning of music remains largely preserved despite their deficient music processing. Implications of the results for statistical learning and intervention programs to mitigate amusia are discussed.
 https://doi.org/10.1121/10.0019472
(Received 6 December 2022; revised 27 March 2023; accepted 3 May 2023; published online 26 May 2023)
[Editor: Andrew Morrison] Pages: 3117–3129

I. INTRODUCTION

Congenital amusia ("amusia" hereafter) is a generic term for musical disabilities (Peretz, 2016). Individuals with amusia ("amusics" hereafter) make up roughly 1.5%–4% of the general population (Peretz and Hyde, 2003; Peretz and Vuvan, 2017). Amusics lack musical abilities common to typical, musically intact individuals. Peretz (2020, pp. 61–62) sums up amusia as follows: "An amusic sings out-of-tune…Amusics can hardly sing without words…They have trouble recognizing familiar tunes…They do not detect wrong notes…," whereas "intentionally singing a false song," such as "O Canada" in the choir, was apparently not an amusic's intention.

A. The negative influence of amusia on pitch processing

The vast literature has documented the negative influence of amusia on music processing, particularly the processing of pitch (e.g., Peretz *et al.*, 2002; Peretz *et al.*, 2003; Peretz *et al.*, 2008). In Ayotte *et al.* (2002), amusics and musically intact listeners were tested to judge whether the musical melodies contained a modified "wrong" note. All the participants were French speakers, and most of them were raised in the francophone culture of Québec. While some amusics reported that one of their parents and certain siblings also had musical problems, members in each family were found not affected, thus eliminating a familial negative attitude toward music as an explanatory factor. Nonetheless, the results of Ayotte *et al.* (2002) revealed that amusics performed close to chance and well below typical listeners. Because of the matched sample demographics, the authors proposed that the observed impairments could not be ascribed to amusics' hearing loss, lack of music exposure, or general cognitive slowing. This musical handicap was broadly and recurrently reported in a plethora of subsequent studies, with results further demonstrating that the deficient pitch processing accounts for the musical disorder and lies at the root of congenital amusia (Ayotte *et al.*, 2002; Foxton *et al.*, 2004; Peretz and Hyde, 2003), either for music perception (Hyde and Peretz, 2003, 2004) or music production (Liu *et al.*, 2013; Tremblay-Champoux *et al.*, 2010). This deficit even extends to pitch-related aspects of speech in a cross-domain manner. For example, Hutchins *et al.* (2010) found that amusics had problems while perceiving fine variations of pitch for intonations. Liu *et al.* (2016) reported that amusics showed poorer performance than typical listeners in lexical tone perception (see also Chen and Peng, 2020; Liu *et al.*, 2021; Ong *et al.*, 2020). The above-mentioned studies indicate that amusia appears to be a domain-general pitch processing disorder (Vuvan *et al.*, 2015).

[a)]Electronic mail: xiaoxiangchensophy@hotmail.com

B. The mental musical lexicon in amusics

It was noted that despite abnormal music processing, amusics were found to possess the mental musical lexicon (Omigie *et al.*, 2012; Omigie *et al.*, 2013; Tillmann *et al.*, 2012). The mental musical lexicon refers to a perceptual representational system for isolated tunes, much in the same way as the mental word lexicon represents isolated words [Peretz *et al.* (2009b), p. 257]. Omigie *et al.* (2012) made use of the priming task to investigate the organization of the mental musical lexicon among amusics and typical listeners. The authors found that amusics were overall slower and less accurate than typical listeners in timbral discrimination of the modified target tone, but they showed a melodic priming effect in that, similar to control listeners, amusics responded faster and more accurately to high-probability than low-probability tones rendered in the same timbre as the context. This was also evident in the study by Tillmann *et al.* (2012), where musical chords, the stimuli used in the current experiment, served as the test materials. Tillmann *et al.* (2012) found that amusics were facilitated in their processing of functionally important in comparison to less important chords in the context of chord sequences, indicating that amusics can develop expectancies for musical events. Furthermore, by adopting a gating paradigm that presents increasingly longer fragments of the stimuli, Tillmann *et al.* (2014) showed that although responding more slowly than did controls, amusics succeeded in differentiating familiar and unfamiliar musical excerpts, although little acoustic information was presented. This pattern was comparable to typical individuals, suggesting that amusics established the mental musical lexicon via their daily exposure to music, despite their deficient music processing (Tillmann *et al.*, 2014).

C. Tracking the statistics from mere exposure by amusics

Previous studies associate this ability of automatic learning from mere exposure with statistical learning, where listeners can track the embedded regularities amid the auditory input and develop the lexicon implicitly without receiving any instruction (Loui, 2022; Ma *et al.*, 2021; Maye *et al.*, 2002; Saffran *et al.*, 1996; Saffran *et al.*, 1999). Nevertheless, findings to date from the amusia literature do not point to a clear role of statistical learning in amusics' establishment of the mental musical lexicon (Loui and Schlaug, 2012; Omigie and Stewart, 2011; Peretz *et al.*, 2012). For example, Peretz *et al.* (2012) played a continuous stream of musical tones to amusic and typical listeners (no silence presented to hint at the boundaries between every tone), with participants required to segment the three-tone motifs based on transitional probability, the conditional likelihood where some event Y will occur given that some other event X has already occurred in the input (Aslin *et al.*, 1998; Saffran *et al.*, 1996). Given the higher transitional probability for the motifs than the non-motifs, typical individuals succeeded in grouping the three-tone motifs versus the non-motifs, yet amusics did not (Peretz *et al.*, 2012). Relatedly, the study by Omigie and Stewart (2011) demonstrated that amusics were less confident in the musical word (motif) segmentation tasks.

Statistical learning has been thought not to rely on transitional probability alone; instead, other statistical cues, such as the stimulus occurrence frequency, also contribute to the subconscious development of the mental lexicon (Erickson and Thiessen, 2015; Maye *et al.*, 2002; Maye *et al.*, 2008; Thiessen, 2017). This possibly relates to the "paradox" in the amusia literature, i.e., between amusics' possession of the mental musical lexicon (Omigie *et al.*, 2012; Omigie *et al.*, 2013; Tillmann *et al.*, 2012; Tillmann *et al.*, 2014) and their partial impairment in tracking the statistics via mere exposure to music, with only the statistical cue of transitional probability examined (Omigie and Stewart, 2011; Peretz *et al.*, 2012). A research gap hence remains to be filled, which concerns whether amusics' mental musical lexicon could be developed based on stimulus occurrence frequency in music exposure. The paradigm of distributional learning (Maye *et al.*, 2002; Maye *et al.*, 2008) provides a valuable opportunity to uncover such a dynamic process, through which amusics' tracking of pitch-related musical regularities, and developing of abstract representations thereof, from the music input can be evaluated.

In one of the seminal studies of distributional learning, Maye *et al.* (2002) used a structured corpus for auditory presentation. The tokens in the corpus were generated from a particular acoustic continuum, and they were always kept the same for all the participants; however, the frequencies that each token would be heard were different, which either constituted a bimodal (tokens near the endpoints occurred most frequently) or a unimodal (tokens in the middle of the continuum occurred most frequently) distribution. Figure 1 schematically displays the two different distributions. In the test after training participants with different distributional structures, learners exposed to the bimodal distribution, rather than those with the unimodal distribution, reliably discriminated the stimuli from the endpoints of the continuum. Evidence of distributional learning has been

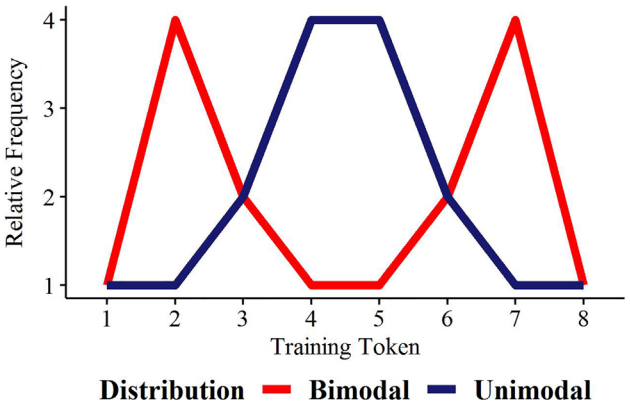

FIG. 1. (Color online) The classic distributional learning paradigm with the bimodal and unimodal distribution conditions based on an eight-step continuum.

https://doi.org/10.1121/10.0019472

preliminarily shown in infants' learning of stop consonants by Maye *et al.* (2002). Later, numerous findings also converged in that the distributional statistical structure of the input influences the building of category membership for both language [consonants (Maye *et al.*, 2002); vowels (Escudero *et al.*, 2011); lexical tones (Ong *et al.*, 2015a,b)] and music [musical chords (Ong *et al.*, 2016; Ong *et al.*, 2017b)] among infant [across the first year (Reh *et al.*, 2021; Werker *et al.*, 2012)] and adult learners [younger adults in their twenties (Ong *et al.*, 2017a) or thirties (Wanrooij *et al.*, 2013); older adults in their seventies (Colby *et al.*, 2018)] over the lifespan. Distributional learning, then, is both a theoretically plausible and experimentally verified method whereby learners learn the sound categories from the auditory input across speech and music domains (Reh *et al.*, 2021). Nonetheless, whether distributional learning would be effective when the auditory tokens are perceived by individuals with impairments, such as congenital amusia, commonly known as "tone deafness" (Peretz and Vuvan, 2017), remains less understood.

D. The present study

In a nutshell, to fully understand amusics' tracking of pitch-related distributional statistics in music exposure, the current study investigated amusics' learnability of music by adopting distributional learning stimuli of musical chord minimal pairs from previous studies (Dean, 2009; Ong *et al.*, 2016; Ong *et al.*, 2017b). Ong *et al.* (2016) and Ong *et al.* (2017b) have shown that normal individuals were able to show distributional learning of a major chord but transposed to a novel microtonal musical scale, the prime number scale (Dean, 2009). This differed from the general musical scale and, thus, diverged from the musical stimuli used in most studies of amusia (e.g., Omigie and Stewart, 2011; Peretz *et al.*, 2012) in which musical melodies based on the chromatic scale were exploited.

Tillmann *et al.* (2012) pointed out that the use of chords (like the use of tones) leads to statistical regularities (differences in frequencies of occurrence and in frequencies of co-occurrence) that can then be internalized by listeners via mere exposure (Bigand and Poulin-Charronnat, 2006). With the minimal pairs of chords based on the prime number scale, we ensured that none of the participants had previous exposure to these stimuli and therefore minimized the influence of participants' pre-existing knowledge on their learning and completion of the tasks [see a similar study by Loui and Schlaug (2012)]. Amusic and control listeners in the current study would be tested twice and compared, following a pretest-training-posttest design (Escudero and Williams, 2014; Maye *et al.*, 2002; Maye *et al.*, 2008). We hypothesized that amusics would show impoverished performance in music perception across tests due to their deficient music processing in the full spectrum (Hyde *et al.*, 2011; Loui *et al.*, 2008; Peretz *et al.*, 2009a). Pertaining to learning outcomes, amusics may exhibit no effects of distributional learning of music; however, it may also be the case that, as implied in prior studies (Omigie *et al.*, 2013; Tillmann *et al.*, 2012; Zhu *et al.*, 2022), perceptual gains in post-training test would be observed in bimodally rather than unimodally trained amusics, akin to normal individuals (Ong *et al.*, 2016; Ong *et al.*, 2017b; Thiessen, 2017).

II. METHOD

A. Participants

Amusic listeners ($n = 18$) and musically intact controls ($n = 19$) participated in the current study. At the beginning, each group consisted of 20 participants, but three of them withdrew from the experiment (one typical and two amusic listeners) due to their unavailability. Participants of the two groups were born and raised in mainland China and self-identified their native language acquired as an infant to be Mandarin Chinese. None of them had lived abroad before the time of testing.

Both amusics and musically intact listeners were identified following the guidelines set by Peretz and colleagues (Peretz and Vuvan, 2017; Peretz *et al.*, 2003; Peretz *et al.*, 2008). First, university students who reported having difficulties experiencing music in everyday life were encouraged to participate in this research project (Peretz *et al.*, 2008; Tillmann *et al.*, 2014). Those with formal musical training were excluded since facilitations in both music and speech processing are frequently reported owing to musicianship (Patel, 2008); also, the exclusion of musicians was beneficial to reach homogeneity among participants. Over 300 individuals were eligible to next be examined via the online Montreal Battery of Evaluation of Amusia (Peretz *et al.*, 2008). This battery with the cutoff score of 71% has been widely used for amusics' identification in previous amusia literature (e.g., Chen and Peng, 2020; Shao and Zhang, 2018; Wang and Peng, 2014; Wong *et al.*, 2012; Zhang *et al.*, 2017), which includes out-of-key, offbeat, and mistuned subtests. The out-of-key and mistuned subtests measure listeners' pitch perception, with the offbeat subtest assessing their rhythm processing (Peretz *et al.*, 2008). Following Peretz and Vuvan (2017), subjects were not allowed to take the test repeatedly. The experimenter monitored the entire test. As revealed by the independent-samples *t*-tests, amusics' global and constituent scores were all significantly lower than the controls ($ps < 0.001$), indicative of their impaired music processing in line with prior studies (Chen and Peng, 2020; Wang and Peng, 2014; Zhang *et al.*, 2017).

The scores obtained by amusics and typical listeners and other background information are exhibited in Table I. It was noted that there were only four male control listeners in the current sample, which was, however, representative of a larger female cohort of undergraduate students of arts in the universities for subject recruitment; meanwhile, gender differences did not appear to have an influence on amusics' musical performance in previous studies (e.g., Ayotte *et al.*, 2002; Peretz *et al.*, 2008). For example, Ayotte *et al.* (2002) proposed that a higher proportion of females was less likely

TABLE I. Demographic characteristics of amusics and controls. The p values indicate results from independent-samples t-tests between amusics and controls.

Subject information	Amusics	Controls	p values
No. of participants (female)	18 (18)	19 (15)	N/A[a]
Age in years (SD)	19.61 (0.85)	19.42 (0.90)	$ps > 0.05$
Working memory (SD)	13.95 (0.97)	14.53 (0.96)	
Online identification test of congenital amusia (SD)			
Out-of-key	63.13 (8.28)	88.72 (7.02)	$ps < 0.001$
Offbeat	60.16 (15.00)	78.69 (8.07)	
Mistuned	56.22 (7.05)	84.17 (8.39)	
Global score	59.83 (6.42)	83.86 (4.59)	

[a]Not applicable (N/A).

related to the condition of congenital amusia but rather reflected the general characteristics of the target participants. Moreover, there was no *a priori* reason to expect gender differences in pitch perception for the present research (Ong *et al.*, 2016). The mean age of each group was 19.61 [standard deviation (SD) = 0.85] for amusics and 19.42 (SD = 0.90) for typical listeners. There was no group difference ($p = 0.51$) in terms of age as suggested by an independent-samples t-test.

Prior to the main experiment, the forward and backward digit span tasks derived from the Wechsler Adult Intelligence Scale-Revised by China (Gong, 1992) were used to measure participants' working memory. This scale has been broadly used in the amusia literature (e.g., Nan *et al.*, 2016; Tang *et al.*, 2018; Yang *et al.*, 2014). There was no group difference in terms of working memory ($p = 0.11$); hence, amusics' general memory capacity was confirmed to be comparable to the controls' in accordance with prior studies (Williamson and Stewart, 2010; Yang *et al.*, 2014; Tillmann *et al.*, 2016). In addition, all the participants were right-handed based on a modified Chinese version of the Edinburgh Handedness Inventory (Oldfield, 1971).

In sum, the two groups in the current study were matched in age, gender, handedness, education, working memory, and musical background. None of the participants reported having hearing loss, brain injuries, or neurological disabilities. All the participants signed the written informed consent and were paid for their participation. The experiment was approved by the ethics review board at the School of Foreign Languages of Hunan University.

B. Materials

The musical materials were all adapted from the preceding studies (Dean, 2009; Ong *et al.*, 2016, 2017b), with two musical chords (chord X and chord Y) synthesized based on a novel microtonal scale of the prime number scale.

The prime number scale has been previously consulted for aesthetic, cognitive, and scientific purposes (Dean, 2009). This novel microtonal scale was developed by first defining a base frequency and next multiplying it with a series of prime numbers (Ong *et al.*, 2017b). Dean (2009) suggests the base frequency ranging from 10 to 20 Hz. The composer is allowed to form a prime number scale that is not "fixed" as the chromatic scale; rather, the scale can be chosen and characterized by the composer. In the current design, the base frequency of 15 Hz was selected with five prime numbers used for multiplication: 17, 23, 31, 41, and 47, following Ong *et al.* (2017b). The reason to start with 17 as the first prime number note (called X1) was that the frequency of the product with the base frequency multiplying the first prime number (i.e., 15 Hz × 17 = 255 Hz) roughly equaled the frequency of middle C in the chromatic scale (C4 at 261.6 Hz). In this respect, the notes of the present prime number scale corresponded to the same scale position as the Western C major scale: as 15 Hz multiplied by 17 referred to C4, the multiplication of the base frequency with the subsequent prime number after 17 (i.e., 15 Hz × 19 = 285 Hz) would refer to the following note after C4 in the C major scale (D4), and so forth. In accordance, the prime number versions of C major (CEG) and G major (GBD) chords were constructed by omitting the prime numbers 19, 29, 37, and 43, which would, respectively, refer to the chromatic notes of D4, F4, A4, and C5. Table II displays the prime number notes with their pitch frequencies. Although any three prime numbers could be chosen to multiply the base frequency to develop the novel chords, certain prime numbers were used, aiming to establish the present prime number scale to correspond to the Western C major scale. The ratios between the tones in the chords from these two scales, however, were different. Figure 2 exhibits comparisons of pitch frequencies of the notes within different musical scales.

Each chord was formed by three musical notes in the current study; namely, chord X consisted of X1, X2, and X3, while chord Y incorporated Y1, Y2, and Y3. A program was written in MaxMSP 5 to specify pitch frequencies of the notes using MIDI, which was then sent to LogicPro 7. Importantly, the MIDI was played using a male choir preset (Choir Male Chant) and a female choir preset (Astral Choir) on the Alchemy plugin (Ong *et al.*, 2017b). All musical stimuli were normalized at 70 dB in intensity. A total of four musical chords were hence synthesized (male chord X, male chord Y, female chord X, and female chord Y), with transposed chord (X and Y) and choir gender (male and female) serving as two test dimensions in the stimuli along which participants would be examined.

TABLE II. Prime number notes exploited in the current experiment, with two chords developed, including chord X (X1-X2-X3) and chord Y (Y1-Y2-Y3). Frequency (in Hz) was displayed for each musical note as below.

Prime number scale	X1	X2	X2′	X3/Y1	Y2	Y2′	Y3
Prime numbers used	17	23	23	31	41	41	47
Prime number notes	255.00	345.00	336.38	465.00	615.00	599.63	705.00

https://doi.org/10.1121/10.0019472

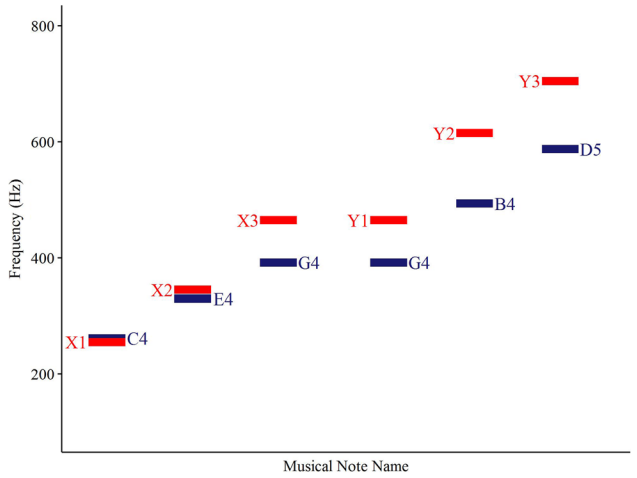

FIG. 2. (Color online) Musical notes in the C major chord (C4-E4-G4) and G major chord (G4-B4-D5) in the Western C major scale and in chord X (X1-X2-X3) and chord Y (Y1-Y2-Y3) in the prime number scale.

The female chord X was used to produce an eight-step pitch continuum for exposure in training, which ensured that the current experiment remained comparable with prior studies (Ong *et al.*, 2016; Ong *et al.*, 2017b). The minimal pair of female chord X was generated through modifying pitch frequency of the middle note (X2) in this chord, with the other two prime number notes (X1 and X3) being constant. Following the study by Koelsch *et al.* (1999), this middle note (X2 at 345 Hz) was 2.5% mistuned to obtain a new one (X2′ at 336.38 Hz), serving as two endpoints along this pitch continuum. The difference between the two middle notes (X2 and X2′) was then calculated and divided by 7 to define the step size in frequency for the intermediate steps (steps 2–7). The X1 and X3 were not interpolated and were sent to LogicPro 7 via MaxMSP 5 and synthesized with the eight middle notes to form the eight chords. Finally, the training continuum was generated, with token 1 being female X1X2X3 and token 8 being female X1X2′X3. Table III documents pitch frequencies of the middle notes gradually shifting from X2 to X2′ in the eight-step pitch continuum. The remaining three minimal pairs, including female chord Y, male chord X, and male chord Y, were synthesized using the same method and would be employed as untrained test chords, aiming to assess whether participants could generalize what they had learned in training (Maye *et al.*, 2008). Each chord was approximately 700 ms in duration; importantly, the duration always remained equivalent within each minimal pair. Five participants who were not involved in the main experiment listened to all these chords and judged them as musical rather than speech materials without any semantic meanings (Ong *et al.*, 2016) and defined them with neutral valence in musical emotion (Bidelman and Walker, 2017). Similar reports also came from an interview that both amusics and controls attended after they finished all tests.

TABLE III. Frequency (in Hz) of X2 as the prime number note for the eight-step musical pitch continuum used in training. The other two prime number notes, X1 and X3, were kept constant in frequency from token 1 to token 8.

Step	1 (X2)	2	3	4	5	6	7	8 (X2′)
Pitch	345.00	343.77	342.54	341.30	340.07	338.84	337.61	336.38

Moreover, there were 32 440-Hz pure tones (70 dB in intensity and 800 ms in duration) synthesized using Praat (Boersma and Weenink, 2018). These tones were randomly interspersed among the training chords and served as the beep tones for a concurrent auditory vigilance task for each participant (Ong *et al.*, 2017b).

C. Procedures

Following previous distributional learning studies (Escudero and Williams, 2014; Maye *et al.*, 2002, 2008), a pretest-training-posttest design was applied into the main experiment. The participants were pseudo-randomly assigned into bimodal or unimodal distribution so that the bimodal group was comparable to the unimodal counterpart (e.g., age, working memory, and test scores in the screening battery). This was beneficial to exclude any potential effects other than stimulus statistical distribution on learning. Four sub-groups were developed, including amusics' bimodal ($n = 9$) and unimodal ($n = 9$) groups and typical listeners' bimodal ($n = 10$) and unimodal ($n = 9$) groups. All the participants listened to the same musical chords in training, yet the occurrences of the tokens differed based on the distribution condition. The stimuli across training and tests were auditorily presented via E-Prime 2.0 (Schneider *et al.*, 2002).

First, the training phase involved mere exposure to the eight tokens generated from the pitch continuum. In the bimodal distribution, two peripheral modal peaks were formed such that tokens 2 and 7 in the training continuum were presented most frequently; nonetheless, in the unimodal distribution, a single central modal peak was developed, with tokens 4 and 5 heard most often. Figure 3 depicts stimulus pitch registers and times of frequencies of the middle note in the chord. It is worth noting that prototypical chords, token 1 (female X1X2X3) and token 8 (female X1X2′X3), were heard with an identical number of times in both conditions. The occurrence frequencies of tokens 1–8 in the bimodal distribution are 16, 64, 32, 16, 16, 32, 64, 16; the occurrence frequencies of tokens 1–8 in the unimodal distribution are 16, 16, 32, 64, 64, 32, 16, 16. The sum of occurrences for the eight training stimuli was 256 in each modal distribution; in addition, 32 beep tones were interspersed within the auditory stream. Thus, a total of 288 sounds were played in a randomized order to each listener. All the participants were required to mark on a paper handout whenever a beep was heard. Their performance in this auditory vigilance task would be checked to ensure that they were indeed engaged in this training phase.

Second, the pretest and posttest shared identical stimuli and common procedures, which were both discrimination

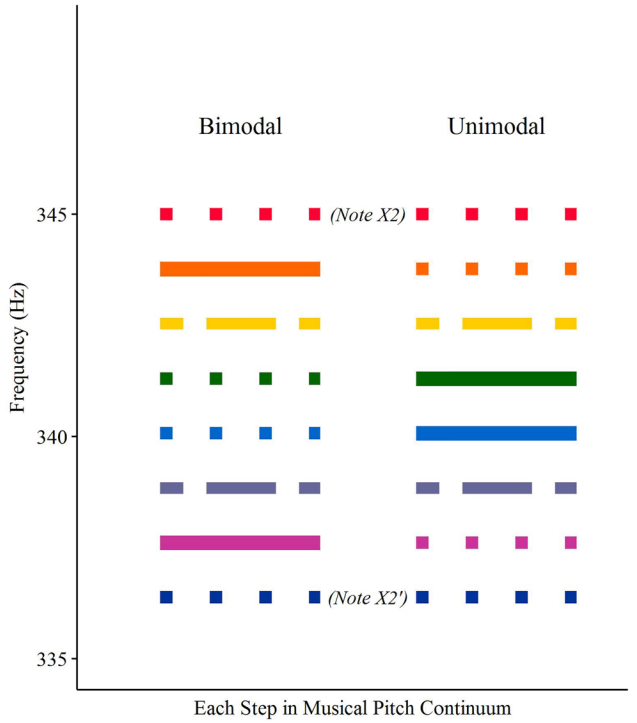

FIG. 3. (Color online) The bimodal and unimodal distributions of musical chords designed in the present study. The eight middle notes are generated from the pitch continuum based on prime number note X2 and its mistuned version X2′. The line types stand for different times of stimulus occurrence. The solid lines (in the bimodal distribution, steps 2 and 7; in the unimodal distribution, steps 4 and 5) represent the stimuli that occur most frequently.

tasks in an ABX format. As an example, after listening to three sounds in a trial, participants were required to decide whether the third sound, X, was similar to the first sound, A, or the second one, B. All four minimal pairs of musical chords, including the untrained ones, were used in discrimination, with each of them presented four times, leading to 32 test trials in both test sessions. The order of trials was randomized. Given amusics' poor pitch processing abilities (Vuvan *et al.*, 2015), participants were not required to respond within a certain time, but they were encouraged to respond as quickly as they could by pressing either of two designated keys on the computer keyboard. In addition, participants were free to take practice trials before the pretest to familiarize themselves with the experimental protocols. The practice materials were not part of the experimental ones, and the feedback was provided only for practice but not the main experiment. The entire distributional learning task, including practice trials, pretest, training, posttest, and a short interview, took around 30–40 min.

D. Data analysis

Accuracy rates were collected and analyzed in the pretest and posttest to examine whether congenital amusics were able to learn pitch-related distributional statistics embedded in music input. Of particular interest was whether perceptual gains from pretest to posttest would be observed in bimodally instead of unimodally trained amusics, as resembling the performance by normal individuals in the distributional learning literature (Erickson and Thiessen, 2015; Maye *et al.*, 2002; Ong *et al.*, 2016; Ong *et al.*, 2017b).

Generalized mixed-effects models were constructed for statistical analysis with the lme4 package (Bates *et al.*, 2015) in R (R Core Team, 2021). The response to each trial was coded as 0 or 1 (incorrect or correct) for each listener. Given two parameters manipulated in stimulus characteristics, responses to transposed chord and choir gender were separately analyzed. In detail, one set of models was built with group (amusics and musically intact individuals), test session (pretest and posttest), modal distribution (bimodal and unimodal), and transposed chord (trained and untrained) acting as fixed factors. The dependent variable was accuracy obtained by participants in both bimodal and unimodal distributions. The other set of similar models was built with group (amusics and musically intact individuals), test session (pretest and posttest), modal distribution (bimodal and unimodal), and choir gender (trained and untrained) acting as fixed factors, with accuracy rates in bimodal and unimodal distributions being the dependent variable. For all models, two-way, three-way, and four-way interaction terms were included as fixed effects. By-subject and by-item random intercepts and slopes for all possible fixed factors were included in the initial model (Barr *et al.*, 2013). When fitting the models, working memory was treated as the controlled covariate. Using the analysis of variance function in lmerTest package (Kuznetsova *et al.*, 2017), the initial model was compared with a simplified model that excluded a specific fixed factor. Pairwise comparisons were calculated with Tukey adjustment using the lsmeans package (Lenth, 2016).

III. RESULTS

In the concurrent auditory vigilance task, the listeners in both groups succeeded in identifying the 32 beeps randomly interspersed in music input; therefore, none of the participants were excluded from further data analysis. Table IV shows the accuracy rates obtained by amusics and typical listeners across pretest and posttest. Amusics' accuracy rates were generally lower than controls; also, pretest scores differed from posttest scores for both amusic and control listeners as a function of the distribution condition.

A. Perceptual accuracy on test dimension of transposed chord

Figure 4 plots accuracy rates by amusic and typical listeners across test session and test dimension of transposed chord as faceted by distribution condition, with error bars

TABLE IV. Mean accuracy rates (SD) of musical chords in the pretest and posttest by bimodally and unimodally trained amusics and musically intact individuals.

	Amusics		Controls	
Test	Bimodal	Unimodal	Bimodal	Unimodal
Pretest	0.78 (0.41)	0.83 (0.37)	0.91 (0.29)	0.94 (0.24)
Posttest	0.89 (0.31)	0.85 (0.36)	0.96 (0.20)	0.95 (0.21)

https://doi.org/10.1121/10.0019472

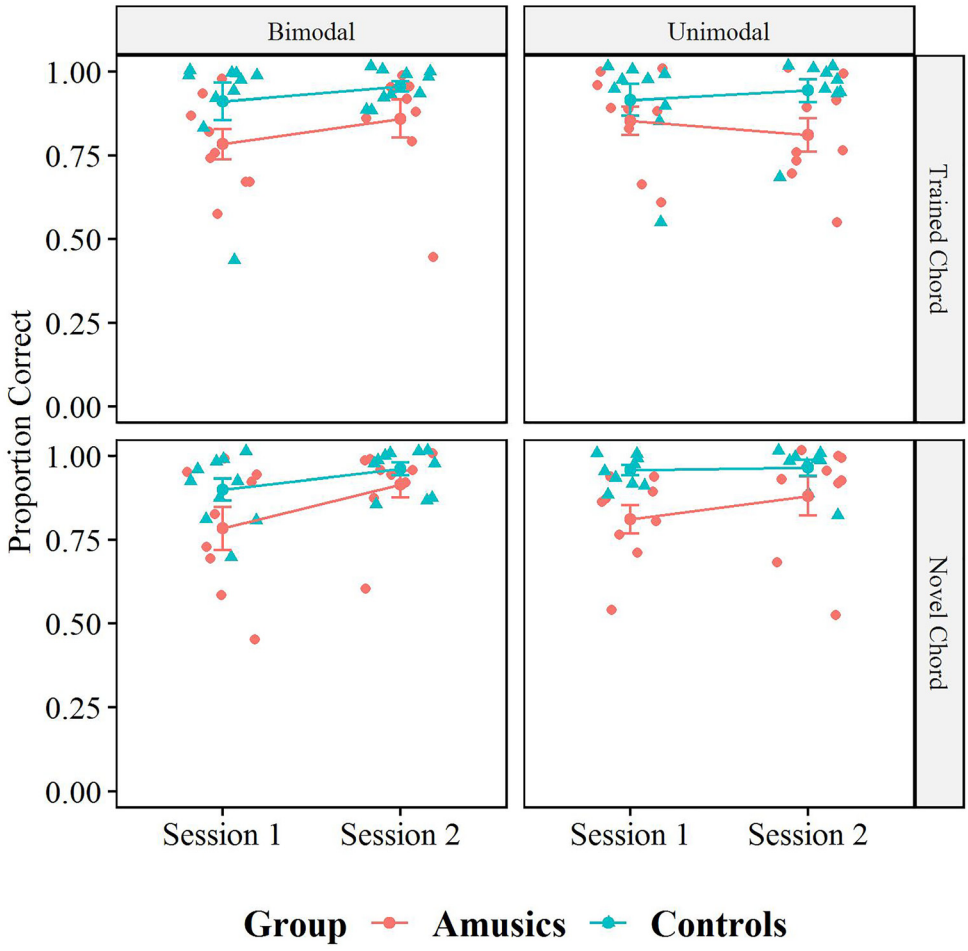

FIG. 4. (Color online) Mean correct responses as measured on the test dimension of transposed chord in different distribution conditions and test sessions by amusics and controls, with error bars showing one SE.

representing one standard error (SE). This figure shows that amusics exhibited a lower accuracy than their counterparts in both tests; moreover, both groups increased their scores in posttest after exposure under the bimodal distribution, whereas this was not obvious between test sessions in the unimodal condition.

The mixed-effects models first revealed a significant main effect of group, χ^2 (1) = 12.14, $p < 0.001$, suggesting that independent of pretest and posttest, amusics [mean (M) = 0.84, SD = 0.37] were outperformed by typical listeners (M = 0.94, SD = 0.24) when listening to musical chords regardless of distribution conditions. There were also a significant main effect of test session, χ^2 (1) = 14.07, $p < 0.001$, and, crucially, a significant two-way interaction between test session and modal distribution, χ^2 (1) = 4.85, $p < 0.05$. Further analysis of this interaction showed that for amusics or controls in the pretest, the scores obtained by the bimodal group did not differ from the unimodal group ($\beta = -0.52$, SE = 0.43, $t = -1.21$, $p = 0.23$). This indicated that the bimodal group showed comparable performance to the unimodal counterpart before training. In addition, in the unimodal distribution, accuracy of pretest did not differ from that of posttest in both amusic and control groups ($\beta = -0.25$, SE = 0.23, $t = -1.06$, $p = 0.29$); in the bimodal distribution, nevertheless, scores obtained by amusics (trained chord: M = 0.86, SD = 0.35; untrained chord: M = 0.92, SD = 0.28) and typical listeners (trained chord: M = 0.96, SD = 0.21; untrained chord: M = 0.96, SD = 0.19) in the posttest were significantly higher than scores obtained by amusics (trained chord: M = 0.78, SD = 0.41; untrained chord: M = 0.78, SD = 0.41) and typical listeners (trained chord: M = 0.91, SD = 0.28; untrained chord: M = 0.90, SD = 0.30) in the pretest ($\beta = -0.97$, SE = 0.22, $t = -4.49$, $p < 0.001$). Other effects did not reach significance ($ps > 0.05$).

In summary, for the test dimension of transposed chord, amusics showed decreased accuracy as compared to musically intact listeners across test sessions, implying their abnormal music processing. However, similar to typical listeners, perceptual gains for both trained and untrained chords were observed in amusics who were bimodally rather than unimodally trained, suggestive of amusics' largely preserved distributional learning of music. Next, statistical analysis was conducted in terms of the test dimension of choir gender.

B. Perceptual accuracy on test dimension of choir gender

Figure 5 plots accuracies by amusic and typical listeners in pretest and posttest, which were faceted by distribution

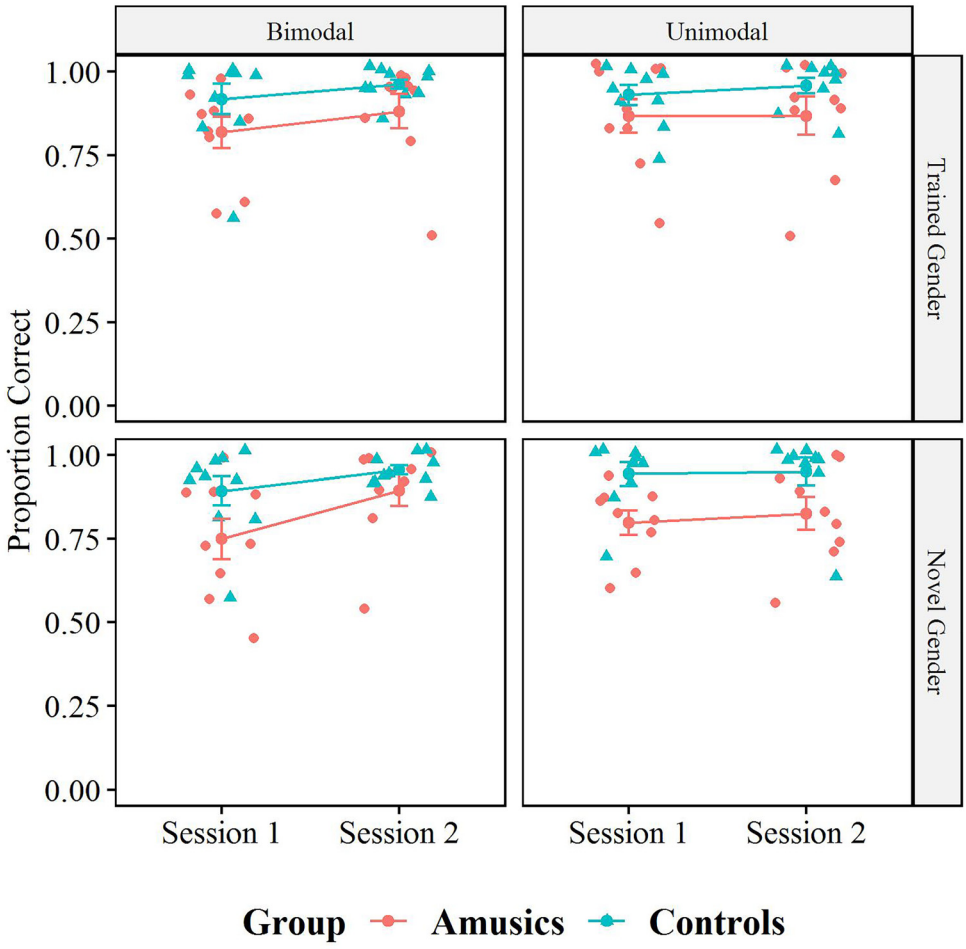

FIG. 5. (Color online) Mean correct responses as measured on the test dimension of choir gender in different distribution conditions and test sessions by amusics and controls, with error bars showing one SE.

condition and test dimension of choir gender with error bars depicting one SE. This figure shows that amusics were graded lower overall as compared to musically intact listeners; in addition, scores from pretest to posttest were improved in the two groups in the bimodal distribution, yet such changes were not seen in the unimodal distribution.

First and foremost, similar to the test dimension of transposed chord, the mixed-effects models uncovered a significant main effect of group, χ^2 (1) = 11.98, $p < 0.001$, for the test dimension of choir gender. This suggested amusics' inferior performance (M = 0.84, SD = 0.37) to typical listeners (M = 0.94, SD = 0.24) when listening to musical chords with either a female (trained) or male (untrained) choir gender across test sessions and distribution conditions. There were a significant main effect of test session, χ^2 (1) = 13.99, $p < 0.001$, and, notably, an interaction between test session and modal distribution, χ^2 (1) = 4.67, $p < 0.05$. Further analysis of this interaction revealed that in the pretest, the scores obtained by the bimodal group did not differ from the unimodal counterpart regardless of amusic or typical listeners ($\beta = -0.48$, SE = 0.43, $t = -1.13$, $p = 0.26$), indicative of the similar performance between the bimodal and unimodal groups before training. In addition, in the unimodal distribution, neither the amusic group nor the control group obtained posttest scores that differed from pretest scores ($\beta = -0.25$, SE = 0.23, $t = -1.08$, $p = 0.28$); however, in the bimodal distribution, amusics (trained gender: M = 0.88, SD = 0.32; untrained gender: M = 0.90, SD = 0.31) and typical listeners (trained gender: M = 0.96, SD = 0.19; untrained gender: M = 0.96, SD = 0.21) were both graded significantly higher in the posttest than in the pretest (amusics: trained gender: M = 0.82, SD = 0.39; untrained gender: M = 0.75, SD = 0.43; controls: trained gender: M = 0.92, SD = 0.27; untrained gender: M = 0.89, SD = 0.31; $\beta = -0.94$, SE = 0.21, $t = -4.39$, $p < 0.001$). Other effects did not reach significance ($ps > 0.05$).

To conclude, for the test dimension of choir gender, there were overall reduced scores obtained by amusics compared with musically intact listeners. The inferior musical performance shown by amusics indicated their impaired musical pitch processing (Peretz *et al.*, 2009a). Notwithstanding, amusics gained improvement in perceptual accuracy after exposure to the bimodal instead of unimodal distribution. This pattern of performance profile was similar to normal individuals in the current and previous studies (Maye *et al.*, 2008; Ong *et al.*, 2017b). Combining the results of the two test dimensions, i.e., transposed chord and choir gender, it was possible to illustrate that amusics'

https://doi.org/10.1121/10.0019472

distributional learning of music was largely spared despite amusia being a musical pitch processing disorder (Peretz and Vuvan, 2017).

IV. DISCUSSION

The current study investigated whether amusics' distributional learning mechanism would function normally when learning musical pitch information embedded in musical chords, which were transposed based on a novel microtonal scale of the prime number scale (Dean, 2009). During the training of mere exposure, amusics and typical listeners were presented with exactly the same musical chords, which only varied in stimulus frequency of statistical distribution, including bimodal and unimodal distributions. Results of perceptual accuracy showed that amusics were outperformed by control listeners across pre-training and post-training tests, suggesting amusics' deficient music processing in the full spectrum; nonetheless, similar to musically intact listeners, amusics who were exposed to the bimodal distribution improved their musical performance, while those in the unimodal condition did not, indicative of amusics' preserved distributional learning of music despite congenital amusia.

A. The deficient music perception in amusics

First, this study corroborated the well-documented finding of amusics' inferior performance to the matched controls in music perception (Ayotte *et al.*, 2002; Peretz *et al.*, 2003; Vuvan *et al.*, 2015). It was repeatedly shown that amusics were outperformed by typical listeners in tasks of either music perception or music production (Peretz, 2016). For example, the study by Hyde and Peretz (2004) revealed that amusics insensitively detected when the fourth tone was displaced in pitch in monotonic sequences of five tones. Likewise, the results of Jiang *et al.* (2010) showed that both identification and discrimination of musical melodies were impaired in amusics relative to typical listeners. Moreover, amusics performed worse than controls in imitation of songs in terms of both absolute and relative pitch matching (Liu *et al.*, 2013), with their singing remaining poor on the pitch dimension although they could benefit from the imitation of songs (i.e., after hearing a model or in unison with the model; Tremblay-Champoux *et al.*, 2010). As amusics were graded lower than typical listeners across test sessions, the current study duplicated amusics' deficient music perception consistently reported in the amusia literature (Foxton *et al.*, 2004; Moreau *et al.*, 2013; Peretz *et al.*, 2022).

A new finding that for the first time complemented previous studies involved amusics' impoverished performance of music perception after distributional learning. Different from many amusia studies using the general musical scale, musical materials in the present study were developed based on a novel microtonal scale of the prime number scale (Dean, 2009). This scale was intentionally exploited to maximally exclude the influence of listeners' pre-existing knowledge on their learning process, which could otherwise contaminate the learning outcome (Loui and Schlaug, 2012). After being exposed to those novel tokens, results of the posttest revealed that amusics, even trained bimodally, obtained lower scores than typical learners, suggesting that congenital amusia persistently interferes with amusics' music processing regardless of modal distribution, test session, and test dimension of stimulus characteristics. While the decreased performance across the board could result from amusics' impaired music processing, what remained more interesting was to scrutinize why bimodally trained amusics still reached a perceptual level lower than musically intact counterparts in the post-training test. This possibly implicated the limited effects of distributional learning of music on amusia or indicated that the musical training programs (i.e., mere exposure) designed in the current study need to be further improved, although this study did not focus on amusics' recovery or intervention. Recently, after primarily reviewing the research on amusia and other developmental disorders, Peretz (2016) pointed out that any training programs to mitigate amusia would be more positive if amusics were provided with external feedback. Future studies may want to design a training program with active feedback given to amusic listeners to, for example, improve their performance in an error-driven manner in distributional learning (Nixon, 2020). In addition, according to Escudero and Williams (2014), distributional learning has far-reaching and long-lasting effects because the improvement by bimodally trained listeners from pretest to the immediate and delayed posttests was maintained after 6 and 12 months. The features of training tokens of vowels, although stemming from the language domain, were purposefully exaggerated in an infant- or foreigner-directed speech style. Because prior studies have shown that abnormal acoustic processing of musical pitch leads to amusics' deficient music perception (Vuvan *et al.*, 2015), future studies may want to employ a modified distributional learning paradigm, i.e., with materials of exaggerated pitch-related properties, to sharpen amusics' sensitivity to stimulus acoustic information in music input.

B. The spared distributional learning of music in amusics

This study provided one more new finding by showing that amusics' performance profiles patterned similarly to typical listeners. That is, bimodally trained amusics obtained higher scores in the posttest than pretest, while the improvement was not found for unimodally trained counterparts (Maye *et al.*, 2002; Maye *et al.*, 2008). Importantly, amusics' improvement in the bimodal distribution was identified across test dimensions of transposed chord and choir gender; namely, regardless of whether the test stimuli were the trained or non-trained tokens, amusics discriminated them better as belonging to different sound clusters, i.e., chord X or chord Y, in the post-training test. This learning outcome accords with generalization effects of distributional learning (e.g., consonants: Maye *et al.*, 2002; Maye *et al.*, 2008; vowels: Escudero *et al.*, 2011; lexical tones: Ong *et al.*,

2015a,b; Ong *et al.*, 2017a; musical chords: Ong *et al.*, 2016; Ong *et al.*, 2017b), which is evident for amusics' preserved distributional learning of music.

This result was not completely unexpected. Although amusia leads to marked difficulties with music processing, Omigie *et al.* (2012) revealed a melodic priming effect among amusics as they responded faster to high-probability versus low-probability musical tones [also see the similar finding plus tone error results in Zhu *et al.* (2022) in the context of linguistic tone]. Based on predictions of a melodic computational model, amusics distinguished tone probabilities in compliance with melodic structure, indicating their sensitivity to musical structure knowledge. Tillmann *et al.* (2012) also disclosed that amusic individuals could develop expectancies for musical events, because they internalized sophisticated syntactic-like functions of musical chords; in addition, Tillmann *et al.* (2014) further unveiled that amusics have built and can access the mental musical lexicon despite amusia. Moreover, a body of evidence from the neurophysiological perspective showed that the amusic brain can automatically track pitch regularities [signaled by frequency-following response (Liu *et al.*, 2015), mismatch negativity (Moreau *et al.*, 2013), and early right anterior negativity (Peretz *et al.*, 2009a)], and there is no anomaly at the level of amusics' cochlear or midbrain, with auditory inputs reaching the superior temporal gyrus normally (Peretz, 2016). All the above-mentioned studies point to the possibility that amusics could obtain higher scores in posttest than pretest via their exposure to music in distributional learning (Ong *et al.*, 2017b; Thiessen, 2017). This sheds some light to the speculation that aberrant processing of a specific cue (e.g., musical pitch) might not preclude distributional learning of that cue. Moreover, because amusics showed impaired computation of transitional probability [as in Peretz *et al.* (2012)] but preserved capacity to track stimulus occurrence frequency (as in our study), this may favor a multicomponent view of statistical learning (Arciuli, 2017; Schneider *et al.*, 2022; Thiessen, 2017). The current study serves as a preliminary effort. Future research may explore these tentative claims more deeply, e.g., with a group of non-tonal-language-speaking amusics (see discussion below).

The results herein, however, diverged from a previous amusia study by Loui and Schlaug (2012). Both the current study and Loui and Schlaug recruited amusic and nonamusic listeners and examined their learning performance after mere exposure to musical stimuli based on the scales different from the existing musical systems, i.e., prime number scale (Dean, 2009) and Bohlen-Pierce musical scale (Loui and Wessel, 2008), respectively. Unlike the current findings, amusics in Loui and Schlaug (2012) only showed significantly impaired learning abilities in pitch-related event frequency information as compared with typical listeners. This discrepancy could possibly be a result of the experimental approach. Loui and Schlaug prepared a corpus consisting of 400 musical melodies, which were not generated from a particular continuum, to be played in exposure; also, their stimulus statistical structure was not adjusted in accordance with distributional learning (Ong *et al.*, 2017a; Ong *et al.*, 2017b). Following the classic distributional learning paradigm (Maye *et al.*, 2002; Maye *et al.*, 2008), the current study, on the contrary, controlled stimulus distributional structure in that frequencies of each token to be heard varied as a function of bimodal and unimodal distribution conditions. As validated in the literature (e.g., Escudero and Williams, 2014; Maye *et al.*, 2002; Maye *et al.*, 2008; Ong *et al.*, 2015a; Ong *et al.*, 2017b), this design could help listeners establish category membership and permitted them to classify the sound identities in the post-training test (Thiessen, 2017), which was evidenced among both trained and untrained test musical chords in our findings.

A general account for the learners' (non-)improvement involved stimulus distributional statistical structure such that the bimodal distribution elicited listeners' perception of the test tokens as exemplars from two discrepant sound categories, but a unimodal distribution induced listeners' perception of the test tokens as exemplars from a single sound category (Maye *et al.*, 2002; Ong *et al.*, 2017b; Thiessen, 2017). To be concrete, participants received mere exposure in the training phase of distributional learning (Escudero and Williams, 2014); that is, they were neither guided on how to memorize the musical or linguistic word nor required to provide any overt responses to each sound. During this process, learners were not simply storing the heard tokens in mind; they were meanwhile unconsciously comparing prior and current exemplars along a certain continuum and then forming an integrated representation (Erickson and Thiessen, 2015). As indicated by the computational model, Integrative Minerva (Thiessen and Pavlik, 2013), the mental representation was gradually developed by computing the variability of exemplars, which were either strengthened or weakened by positive or negative feature valence in the stimuli. The updated mental lexicon would be eventually obtained, which mirrored the central tendency of the distributions. Although amusics were indeed impaired when processing musical pitch, they have the capacity to track pitch regularities subliminally (e.g., Liu *et al.*, 2015; Moreau *et al.*, 2013; Peretz *et al.*, 2009a). This supported that amusics' mental musical lexicon could be developed from mere exposure as controls, regardless of their deficient musical pitch processing. One more piece of direct evidence came from a music gating experiment, which demonstrated that amusics could build and access the mental musical lexicon with accumulated everyday experience of music, although insufficient acoustic information in musical excerpts was heard (Tillmann *et al.*, 2014).

There was one more possible explanation that concerned the musical stimuli used in the current experiment. Many previous studies examined amusics' music perception by adopting melodic materials that unfold musical tones with time (e.g., Omigie and Stewart, 2011; Peretz *et al.*, 2012). While melodies are monophonic tone sequences (i.e., only one tone is played at a time), harmonic structures are based on chords and add a vertical dimension to melodic

https://doi.org/10.1121/10.0019472

structures, with chords being created by three or four tones played simultaneously (Tillmann *et al.*, 2012). Musical chords in this study were developed by three tones (chord X by X1, X2, and X3; chord Y by Y1, Y2, and Y3) based on the prime number scale (Dean, 2009). In line with the priming study by Tillmann *et al.* (2012), our results duplicated as well as reinforced the view that amusics were able to learn musical structure, i.e., the regularities of associations between pitches, which was even constructed according to a novel microtonal scale that none of the listeners had known before. This also confirmed that amusics show a sensitivity to musical structure as previously measured either behaviorally (Omigie *et al.*, 2012) or neuropsychologically (Omigie *et al.*, 2013). It is worth noting that although amusics could involuntarily establish pitch-related representations as controls, their behavioral performance shown in the pretest and posttest remained degraded, possibly due to amusia as a disconnection syndrome that impairs conscious usage of musical knowledge from higher-level (the inferior frontal gyrus) cortical parts onto lower-level (the superior temporal gyrus) auditory processes (Hyde *et al.*, 2011; Loui *et al.*, 2009; Peretz, 2016).

We also noticed a potential explanation with respect to the participants' language background. All the participants were native speakers of Mandarin Chinese as a tone language. Previous studies have shown that tonal-language speakers have musical competence superior to that of non-tonal-language speakers. For example, it was found that as compared with non-tonal-language speakers, tonal-language speakers were more likely to have absolute-pitch labeling abilities and were better able to perceptually discriminate musical pitch (Deutsch *et al.*, 2004; Deutsch *et al.*, 2006; Pfordresher and Brown, 2009). In addition, Wong *et al.* (2012) conducted a large-scale study in Cantonese speakers and calculated a lower prevalence of amusia than previously reported in Western populations. They additionally revealed that tonal-language-speaking amusics outperformed non-tonal-language-speaking amusics in musical pitch perception. It was, hence, less clear whether our results were partially attributable to amusics' tone language experience. In this regard, future studies may want to recruit non-tonal-language speakers with amusia and examine their distributional learning performance as compared with the present study. This would help clarify whether the current findings could be generalized onto non-tonal-language speakers with amusia and whether the language background contributed to the differences between our findings and those of Peretz *et al.* (2012).

Last but not least, amusics' largely preserved distributional learning of music is informative for the treatment of amusia, since it indicates that amusics are able to establish, sustain, and access the abstract pitch-related mental categories (Thiessen, 2017). This does not deny the fact that amusics are not equally musical and are born that way as compared to the general population (Peretz *et al.*, 2022). In addition, future studies may want to evaluate whether the distributional learning paradigm could strengthen the mental lexicon for other clinical populations who were reported as having weakened phonological representations or lexical access deficits [e.g., individuals who stutter (Howell and Bernstein Ratner, 2018) and individuals with aphasia (Mirman and Britt, 2014)]. These endeavors of intervention are of high significance both theoretically and practically in rehabilitating people with developmental disorders and understanding the relationship between statistical learning and these deficits.

C. Limitations and future directions

We conclude by pointing out the limitations of the current study. First, although prior studies have not reported any gender differences in amusics' music perception (e.g., Ayotte *et al.*, 2002; Peretz *et al.*, 2008), our groups consisted of a large cohort of female participants. While this was representative of the sample available for subject recruitment and might not be a reason for group differences in test performance, replicating the study with a sample that includes more female participants or includes participants in a more typical gender ratio could strengthen the findings of this study. Second, although the current musical tokens were synthesized in different dimensions, the introduction of stimuli with more variability would make the laboratory setting more ecological. Third, because amusia is a domain-general pitch processing disorder, it would be interesting to further evaluate amusics' performance in distributional learning of non-native lexical tones. Meanwhile, it is suggested that future studies evaluate amusics' cerebral correlates of distributional learning with the aid of neuroimaging tools. The comparison of performance profiles between amusics and controls as measured at both behavioral and neurobiological levels would advance the understanding not only of the pathology of congenital amusia but also of intervention programs to mitigate amusia (Peretz, 2016).

ACKNOWLEDGMENTS

This study was supported by grants from the Humanities and Social Science Project of Ministry of Education of China (Grant Nos. 22YJC740008 and 22YJC740093) and the Fundamental Research Funds for the Central Universities of China (Grant No. 531118010660). This study involving human participants was reviewed and approved by the ethics review board at the School of Foreign Languages of Hunan University. Written informed consent was signed by each participant. The authors have declared that no competing interests existed at the time of publication. The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Arciuli, J. (**2017**). "The multi-component nature of statistical learning," Philos. Trans. R. Soc. B **372**(1711), 20160058.

Aslin, R. N., Saffran, J. R., and Newport, E. L. (**1998**). "Computation of conditional probability statistics by 8-month-old infants," Psychol. Sci. **9**(4), 321–324.

Ayotte, J., Peretz, I., and Hyde, K. (**2002**). "Congenital amusia: A group study of adults afflicted with a music-specific disorder," Brain **125**(2), 238–251.
Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (**2013**). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," J. Mem. Lang. **68**(3), 255–278.
Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.
Bidelman, G. M., and Walker, B. S. (**2017**). "Attentional modulation and domain-specificity underlying the neural organization of auditory categorical perception," Eur. J. Neurosci. **45**(5), 690–699.
Bigand, E., and Poulin-Charronnat, B. (**2006**). "Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training," Cognition **100**(1), 100–130.
Boersma, P., and Weenink, D. (**2018**). "Praat: Doing phonetics by computer (version 6.0.37) [computer program]," http://www.praat.org (Last viewed 21 December 2018).
Chen, F., and Peng, G. (**2020**). "Reduced sensitivity to between-category information but preserved categorical perception of lexical tones in tone language speakers with congenital amusia," Front. Psychol. **11**, 581410.
Colby, S. E., Clayards, M., and Baum, S. R. (**2018**). "The role of lexical status and individual differences for perceptual learning in younger and older adults," J. Speech Lang. Hear. Res. **61**(8), 1855–1874.
Dean, R. T. (**2009**). "Widening unequal tempered microtonal pitch space for metaphoric and cognitive purposes with new prime number scales," Leonardo **42**(1), 94–95.
Deutsch, D., Henthorn, T., and Dolson, M. (**2004**). "Absolute pitch, speech, and tone language: Some experiments and a proposed framework," Music Percept. **21**(3), 339–356.
Deutsch, D., Henthorn, T., Marvin, E., and Xu, H. (**2006**). "Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period," J. Acoust. Soc. Am. **119**(2), 719–722.
Erickson, L. C., and Thiessen, E. D. (**2015**). "Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition," Dev. Rev. **37**, 66–108.
Escudero, P., Benders, T., and Wanrooij, K. (**2011**). "Enhanced bimodal distributions facilitate the learning of second language vowels," J. Acoust. Soc. Am. **130**(4), EL206–EL212.
Escudero, P., and Williams, D. (**2014**). "Distributional learning has immediate and long-lasting effects," Cognition **133**(2), 408–413.
Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., and Griffths, T. D. (**2004**). "Characterization of deficits in pitch perception underlying 'tone deafness,' " Brain **127**(4), 801–810.
Gong, Y. X. (**1992**). *Wechsler Adult Intelligence Scale-Revised in China Version* (Hunan Medical College, Changsha, China).
Howell, T. A., and Bernstein Ratner, N. (**2018**). "Use of a phoneme monitoring task to examine lexical access in adults who do and do not stutter," J. Fluency Disord. **57**, 65–73.
Hutchins, S., Gosselin, N., and Peretz, I. (**2010**). "Identification of changes along a continuum of speech intonation is impaired in congenital amusia," Front. Psychol. **1**, 236.
Hyde, K. L., and Peretz, I. (**2003**). "'Out-of-pitch' but still 'in-time,' " Ann. N.Y. Acad. Sci. **999**(1), 173–176.
Hyde, K. L., and Peretz, I. (**2004**). "Brains that are out of tune but in time," Psychol. Sci. **15**(5), 356–360.
Hyde, K. L., Zatorre, R. J., and Peretz, I. (**2011**). "Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia," Cereb. Cortex **21**(2), 292–299.
Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., and Yang, Y. (**2010**). "Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese," Neuropsychologia **48**(9), 2630–2639.
Koelsch, S., Schröger, E., and Tervaniemi, M. (**1999**). "Superior preattentive auditory processing in musicians," Neuroreport **10**(6), 1309–1313.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2017**). "LmerTest Package: Tests in linear mixed effects models," J. Stat. Softw. **82**(13), 1–26.
Lenth, R. V. (**2016**). "Least-squares means: The R package lsmeans," J. Stat. Softw. **69**(1), 1–33.
Liu, F., Chan, A. H., Ciocca, V., Roquet, C., Peretz, I., and Wong, P. C. (**2016**). "Pitch perception and production in congenital amusia: Evidence from Cantonese speakers," J. Acoust. Soc. Am. **140**(1), 563–575.
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., Maggu, A. R., Lau, J. C. Y., and Wong, P. C. (**2015**). "Brainstem encoding of speech and musical stimuli in congenital amusia: Evidence from Cantonese speakers," Front. Hum. Neurosci. **8**, 1029.
Liu, F., Yin, Y., Chan, A., Yip, V., and Wong, P. C. (**2021**). "Individuals with congenital amusia do not show context-dependent perception of tonal categories," Brain Lang. **215**(1), 104908.
Loui, P. (**2022**). "New music system reveals spectral contribution to statistical learning," Cognition **224**, 105071.
Loui, P., Alsop, D., and Schlaug, G. (**2009**). "Tone deafness: A new disconnection syndrome?," J. Neurosci. **29**(33), 10215–10220.
Loui, P., Guenther, F. H., Mathys, C., and Schlaug, G. (**2008**). "Action-perception mismatch in tone-deafness," Curr. Biol. **18**(8), R331–R332.
Loui, P., and Schlaug, G. (**2012**). "Impaired learning of event frequencies in tone deafness," Ann. N.Y. Acad. Sci. **1252**(1), 354–360.
Loui, P., and Wessel, D. (**2008**). "Learning and liking an artificial musical system: Effects of set size and repeated exposure," Music Sci. **12**(2), 207–230.
Ma, J., Zhu, J., Yang, Y., and Chen, F. (**2021**). "The development of categorical perception of segments and suprasegments in Mandarin-speaking preschoolers," Front. Psychol. **12**, 693366.
Maye, J., Weiss, D. J., and Aslin, R. N. (**2008**). "Statistical phonetic learning in infants: Facilitation and feature generalization," Dev. Sci. **11**(1), 122–134.
Maye, J., Werker, J. F., and Gerken, L. (**2002**). "Infant sensitivity to distributional information can affect phonetic discrimination," Cognition **82**(3), B101–B111.
Mirman, D., and Britt, A. E. (**2014**). "What we talk about when we talk about access deficits," Philos. Trans. R. Soc. B **369**, 20120388.
Moreau, P., Jolicoeur, P., and Peretz, I. (**2013**). "Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials," Brain Cogn. **81**(3), 337–344.
Nan, Y., Huang, W. T., Wang, W. J., Liu, C., and Dong, Q. (**2016**). "Subgroup differences in the lexical tone mismatch negativity (MMN) among Mandarin speakers with congenital amusia," Biol. Psychol. **113**, 59–67.
Nixon, J. S. (**2020**). "Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking," Cognition **197**, 104081.
Oldfield, R. C. (**1971**). "The assessment and analysis of handedness: The Edinburgh inventory," Neuropsychologia **9**(1), 97–113.
Omigie, D., Pearce, M. T., and Stewart, L. (**2012**). "Tracking of pitch probabilities in congenital amusia," Neuropsychologia **50**(7), 1483–1493.
Omigie, D., Pearce, M. T., Williamson, V. J., and Stewart, L. (**2013**). "Electrophysiological correlates of melodic processing in congenital amusia," Neuropsychologia **51**(9), 1749–1762.
Omigie, D., and Stewart, L. (**2011**). "Preserved statistical learning of tonal and linguistic material in congenital amusia," Front. Psychol. **2**, 109.
Ong, J. H., Burnham, D., and Escudero, P. (**2015a**). "Distributional learning of lexical tones: A comparison of attended vs. unattended listening," PLoS One **10**(7), e0133446.
Ong, J. H., Burnham, D., and Escudero, P. (**2015b**). "Mandarin listeners can learn non-native lexical tones through distributional learning," in *Proceedings of the 18th International Congress of Phonetic Sciences*, August 10–14, Glasgow, UK.
Ong, J. H., Burnham, D., Escudero, P., and Stevens, C. J. (**2017a**). "Effect of linguistic and musical experience on distributional learning of nonnative lexical tones," J. Speech Lang. Hear. Res. **60**(10), 2769–2780.
Ong, J. H., Burnham, D., and Stevens, C. J. (**2017b**). "Learning novel musical pitch via distributional learning," J. Exp. Psychol. Learn. Mem. Cogn. **43**(1), 150–157.
Ong, J. H., Burnham, D., Stevens, C. J., and Escudero, P. (**2016**). "Naive learners show cross-domain transfer after distributional learning: The case of lexical and musical pitch," Front. Psychol. **7**, 1189.
Ong, J. H., Tan, S. H., Chan, A. H. D., and Wong, F. C. K. (**2020**). "The effect of musical experience and disorders on lexical tone perception,

production, and learning: A review," in *Speech Learning, Perception, and Production: Multidisciplinary Approaches in Chinese Language Research*, edited by M. Liu, F.-M. Tsao, and P. Li (Springer Nature, Singapore), pp. 139–158.
Patel, A. D. (**2008**). *Music, Language, and the Brain* (Oxford University, New York).
Peretz, I. (**2016**). "Neurobiology of congenital amusia," Trends Cogn. Sci. **20**(11), 857–867.
Peretz, I. (**2020**). *How Music Sculpts Our Brain* (Odile Jacob, Paris).
Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., and Jutras, B. (**2002**). "Congenital amusia: A disorder of fine-grained pitch discrimination," Neuron **33**(2), 185–191.
Peretz, I., Brattico, E., Järvenpää, M., and Tervaniemi, M. (**2009a**). "The amusic brain: In tune, out of key, and unaware," Brain **132**(5), 1277–1286.
Peretz, I., Champod, A. S., and Hyde, K. (**2003**). "Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia," Ann. N.Y. Acad. Sci. **999**(1), 58–75.
Peretz, I., Gosselin, N., Belin, P., Zatorre, R. J., Plailly, J., and Tillmann, B. (**2009b**). "Music lexical networks: The cortical organization of music recognition," Ann. N.Y. Acad. Sci. **1169**(1), 256–265.
Peretz, I., Gosselin, N., Tillmann, B., Cuddy, L. L., Gagnon, B., Trimmer, G. C., Paquette, S., and Bouchard, B. (**2008**). "On-line identification of congenital amusia," Music Percept. **25**(4), 331–343.
Peretz, I., and Hyde, K. L. (**2003**). "What is specific to music processing? Insights from congenital amusia," Trends Cogn. Sci. **7**(8), 362–367.
Peretz, I., Ross, J., Bourassa, C. V., Perreault, L.-P. L., Dion, P. A., Weiss, M. W., Felezeu, M., Rouleau, G. A., and Dubé, M.-P. (**2022**). "Do variants in the coding regions of FOXP2, a gene implicated in speech disorder, confer a risk for congenital amusia?," Ann. N.Y. Acad. Sci. **1517**(1), 275–285.
Peretz, I., Saffran, J., Schön, D., and Gosselin, N. (**2012**). "Statistical learning of speech, not music, in congenital amusia," Ann. N.Y. Acad. Sci. **1252**(1), 361–367.
Peretz, I., and Vuvan, D. T. (**2017**). "Prevalence of congenital amusia," Eur. J. Hum. Genet. **25**(5), 625–630.
Pfordresher, P. Q., and Brown, S. (**2009**). "Enhanced production and perception of musical pitch in tone language speakers," Atten. Percept. Psychophys. **71**, 1385–1398.
R Core Team (**2021**). "R: A language and environment for statistical computing," http://www.R-project.org/ (Last viewed 27 December 2021).
Reh, R. K., Hensch, T. K., and Werker, J. F. (**2021**). "Distributional learning of speech sound categories is gated by sensitive periods," Cognition **213**, 104653.
Saffran, J. R., Aslin, R. N., and Newport, E. L. (**1996**). "Statistical learning by 8-month-old infants," Science **274**(5294), 1926–1928.
Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (**1999**). "Statistical learning of tone sequences by human infants and adults," Cognition **70**(1), 27–52.
Schneider, J., Weng, Y. L., Hu, A., and Qi, Z. (**2022**). "Linking the neural basis of distributional statistical learning with transitional statistical learning: The paradox of attention," Neuropsychologia **172**, 108284.
Schneider, W., Eschman, A., and Zuccolotto, A. (**2002**). *E-Prime: User's Guide* (Psychology Software, Pittsburgh, PA).
Shao, J., and Zhang, C. (**2018**). "Context integration deficit in tone perception in Cantonese speakers with congenital amusia," J. Acoust. Soc. Am. **144**(4), EL333–EL339.
Tang, W., Wang, X. J., Li, J. Q., Liu, C., Dong, Q., and Nan, Y. (**2018**). "Vowel and tone recognition in quiet and in noise among Mandarin-speaking amusics," Hear. Res. **363**, 62–69.
Thiessen, E. D. (**2017**). "What's statistical about learning? Insights from modelling statistical learning as a set of memory processes," Phil. Trans. R. Soc. B **372**(1711), 20160056.
Thiessen, E. D., and Pavlik, P. I. (**2013**). "iMinerva: A mathematical model of distributional statistical learning," Cogn. Sci. **37**(2), 310–343.
Tillmann, B., Albouy, P., Caclin, A., and Bigand, E. (**2014**). "Musical familiarity in congenital amusia: Evidence from a gating paradigm," Cortex **59**, 84–94.
Tillmann, B., Gosselin, N., Bigand, E., and Peretz, I. (**2012**). "Priming paradigm reveals harmonic structure processing in congenital amusia," Cortex **48**(8), 1073–1078.
Tillmann, B., Lévêque, Y., Fornoni, L., Albouy, P., and Caclin, A. (**2016**). "Impaired short-term memory for pitch in congenital amusia," Brain Res. **1640**(Part B), 251–263.
Tremblay-Champoux, A., Dalla Bella, S., Phillips-Silver, J., Lebrun, M.-A., and Peretz, I. (**2010**). "Singing proficiency in congenital amusia: Imitation helps," Cogn. Neuropsychol. **27**(6), 463–476.
Vuvan, D. T., Nunes-Silva, M., and Peretz, I. (**2015**). "Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia," Cortex **69**, 186–200.
Wang, X., and Peng, G. (**2014**). "Phonological processing in Mandarin speakers with congenital amusia," J. Acoust. Soc. Am. **136**(6), 3360–3370.
Wanrooij, K., Escudero, P., and Raijmakers, M. E. J. (**2013**). "What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning," J. Phon. **41**(5), 307–319.
Werker, J. F., Yeung, H. H., and Yoshida, K. A. (**2012**). "How do infants become experts at native-speech perception?," Curr. Dir. Psychol. Sci. **21**(4), 221–226.
Williamson, V. J., and Stewart, L. (**2010**). "Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem," Memory **18**(6), 657–669.
Wong, P. C., Ciocca, V., Chan, A. H., Ha, L. Y., Tan, L. H., and Peretz, I. (**2012**). "Effects of culture on musical pitch perception," PLoS One **7**(4), e33424.
Yang, W. X., Feng, J., Huang, W. T., Zhang, C. X., and Nan, Y. (**2014**). "Perceptual pitch deficits coexist with pitch production difficulties in music but not Mandarin speech," Front. Psychol. **4**, 1024.
Zhang, C., Peng, G., Shao, J., and Wang, W. S. Y. (**2017**). "Neural bases of congenital amusia in tonal language speakers," Neuropsychologia **97**, 18–28.
Zhu, J., Chen, X., Chen, F., and Wiener, S. (**2022**). "Individuals with congenital amusia show degraded speech perception but preserved statistical learning for tone languages," J. Speech Lang. Hear. Res. **65**(1), 53–69.