Neuropsychologia 66 (2015) 111–118

Contents lists available at ScienceDirect

Neuropsychologia

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

A music perception disorder (congenital amusia) influences speech comprehension

Fang Liu [a], Cunmei Jiang [b], Bei Wang [c], Yi Xu [a], Aniruddh D. Patel [d,*]

[a] Department of Speech, Hearing and Phonetic Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF, UK
[b] Music College, Shanghai Normal University, 100 East Guilin Road, Shanghai 200234, China
[c] Institute of Chinese Minority Language and Classics, Minzu University of China, 27 South Zhongguancun Street, Beijing 100081, China
[d] Department of Psychology, Tufts University, 490 Boston Avenue, Medford, MA 02155, USA

ARTICLE INFO

Article history:
Received 20 May 2014
Received in revised form
27 October 2014
Accepted 4 November 2014
Available online 8 November 2014

Keywords:
Congenital amusia
Speech comprehension
Music perception
Pitch
Segment
Noise

ABSTRACT

This study investigated the underlying link between speech and music by examining whether and to what extent congenital amusia, a musical disorder characterized by degraded pitch processing, would impact spoken sentence comprehension for speakers of Mandarin, a tone language. Sixteen Mandarin-speaking amusics and 16 matched controls were tested on the intelligibility of news-like Mandarin sentences with natural and flat fundamental frequency (F_0) contours (created via speech resynthesis) under four signal-to-noise (SNR) conditions (no noise, +5, 0, and −5 dB SNR). While speech intelligibility in quiet and extremely noisy conditions (SNR = −5 dB) was not significantly compromised by flattened F_0, both amusic and control groups achieved better performance with natural-F_0 sentences than flat-F_0 sentences under moderately noisy conditions (SNR = +5 and 0 dB). Relative to normal listeners, amusics demonstrated reduced speech intelligibility in both quiet and noise, regardless of whether the F_0 contours of the sentences were natural or flattened. This deficit in speech intelligibility was not associated with impaired pitch perception in amusia. These findings provide evidence for impaired speech comprehension in congenital amusia, suggesting that the deficit of amusics extends beyond pitch processing and includes segmental processing.

1. Introduction

Music and language are fundamental traits of human existence. It has been suggested that the two are intricately linked in terms of evolution and cognitive processing (e.g., Darwin, 1871; Patel, 2008; Arbib, 2013). In particular, there has been increasing evidence that pitch processing in music and speech shares cognitive and neural mechanisms, with one line of evidence coming from congenital amusia (amusia hereafter), a neurogenetic disorder of musical perception and production (Drayna et al., 2001; Peretz et al., 2007). It has been estimated that amusia affects around 4% of the general population for speakers of both tone and non-tonal languages (Kalmus and Fry, 1980; Nan et al., 2010, though cf. Henry and McAuley, 2010). Individuals with amusia (amusics hereafter) have difficulty detecting out-of-tune notes in melodies and cannot sing in tune despite having normal hearing, intelligence, and regular exposure to music (Ayotte et al., 2002). Apart from musical problems, amusics also demonstrate elevated thresholds for pitch change detection and pitch direction discrimination (Foxton et al., 2004; Hyde and Peretz, 2004; Jiang et al., 2013; Liu et al., 2012b) and impaired memory for pitch (Albouy et al., 2013; Tillmann et al., 2009; Williamson and Stewart, 2010).

Recent research suggests that amusia is not a music-specific disorder, as amusics exhibit impaired performance on lexical tone perception, linguistic and emotional prosody processing, phonological processing, and speech imitation in laboratory conditions (Jiang et al., 2012b; Jones et al., 2009a; Liu et al., 2013, 2010; Liu and Jiang et al., 2012; Nan et al., 2010; Patel et al., 2008; Thompson et al., 2012). Amusics also showed impaired judgment of semantic acceptability of sentences during speech comprehension, and failed to show a P600 response to a mismatch between prosody and syntax (Jiang et al., 2012a). However, the experimental conditions in the aforementioned studies, which involved isolated words or phrases specially constructed for laboratory studies, almost never occur in everyday life. Whether amusia impacts speech comprehension in everyday listening situations remains an open question.

The comprehension of spoken language relies at least partially on prosodic decoding of lexical, syntactic, and discourse

* Corresponding author. Fax: +1 617 627 3181.
E-mail address: a.patel@tufts.edu (A.D. Patel).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.11.001
0028-3932/

information (Cutler et al., 1997). In tone languages like Mandarin, lexical tones are used to distinguish word meaning at the syllable level. For example, /ma/ means "mother" with a high tone, "hemp" with a rising tone, "horse" with a low tone, and "to scold" with a falling tone. It has been found that tones in Mandarin carry functional loads as high as those of vowels in distinguishing words (Surendran and Levow, 2004). However, speech intelligibility in a quiet setting is not necessarily affected by synthetically-flattened fundamental frequency (F_0) for either tone-language or non-tonal language listeners (Binns and Culling, 2007; Patel et al., 2010; Xu et al., 2013). This may be because, in the absence of F_0 variations, syntactic/semantic information as well as non-pitch related aspects of prosody can help speech understanding in quiet. In other words, performance may be at ceiling for speech intelligibility in quiet so that the effect of F_0 manipulation (flat versus natural) cannot be observed. Nevertheless, speech is understood better with natural than flattened F_0 in noisy conditions for listeners of both tone and non-tonal .language backgrounds (Miller et al., 2010; Patel et al., 2010), which may be an indication that natural F_0 contours carry critical information that is especially useful for speech comprehension in noise.

It remains to be determined whether amusia, a deficit in pitch processing, impacts speech comprehension in everyday listening situations. The current study investigated this issue by examining Mandarin-speaking amusics' and controls' comprehension of natural- versus flat-F_0 sentences in quiet and noise. We also examined whether speech comprehension was associated with musical perception abilities and pitch direction identification thresholds for each group. The use of Mandarin speakers allowed us to examine the impact of F_0 manipulation (natural versus flat F_0) on speech comprehension in amusics who speak a tone language, where each syllable is associated with a specific tone. The use of noisy speech was to simulate real-life listening situations.

We predicted that amusia would influence speech comprehension in quiet, as amusics have shown impaired processing of speech sounds in quiet, including word discrimination, syllable segmentation, and phonological processing (Jones et al., 2009a). Recent findings of amusics' impairments in the processing of resolved harmonics (Cousineau et al., 2013) and musical harmonicity (Cousineau et al., 2012) would also predict impaired speech comprehension in noise in amusia, since, similar to harmonicity processing, hearing speech in noise requires concurrent sound segregation (Schwartz et al., 2003). Furthermore, amusics' deficits in pitch processing would be detrimental to speech comprehension especially when pitch plays an essential role in signifying sentence meaning (Jiang et al., 2010, 2012a; Liu et al., 2010). Thus, the combined deficits in segmental, harmonicity, and pitch processing in amusia are likely to have consequences for speech comprehension in quiet and noise, both with natural and flattened F_0 contours.

2. Materials and methods

2.1. Participants

Sixteen amusics and 16 matched controls were recruited through advertisements in the bulletin board system of universities in Shanghai, China. All participants spoke Mandarin Chinese as their native language. The presence or absence of amusia in these participants was assessed using the Montreal Battery of Evaluation of Amusia (MBEA) (Peretz et al., 2003), which consists of six subtests measuring scale, contour, interval, rhythm, meter, and memory processing of melodies. Those who scored below 72% correct on the MBEA global score, which corresponds to two standard deviations below normals' mean, were classified as amusic (Nan et al., 2010). At the time of testing, participants were all enrolled as undergraduate or postgraduate students at universities in Shanghai. None of the participants reported having learning or memory problems, neurological or psychiatric disorders, or speech and hearing difficulties in questionnaires regarding their musical, linguistic, biological, and medical background. Out of the 32 participants, only one amusic received extracurricular musical training for half a year during childhood. None of the other participants received any formal musical training. All but one control participant were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Ethical approval was granted by Shanghai Normal University in China. Written informed consents were obtained from all participants prior to the experiment.

Table 1 shows the characteristics of the amusic and control participants (see Supplementary Table 1 for detailed information). The two groups were matched in sex, handedness, age, and years of education. The amusics performed significantly worse than the controls on all MBEA subtests. All participants also completed a set of four psychophysical pitch threshold tasks, during which they were required to identify the direction of pitch movement (high-low versus low-high; rising-falling versus falling-rising) in the

Table 1
Characteristics of the amusic (n=16) and control (n=16) groups.

Group	Sex	Handedness	Age	Education	Scale	Contour	Interval	Rhythm	Meter	Memory	MBEA pitch composite	Pitch direction threshold
Control												
Mean	10 F	1 L	23.06	16.88	28.00	28.19	28.44	28.31	24.06	26.44	84.63	0.40
(SD)	6 M	15 R	(1.18)	(0.81)	(1.63)	(1.33)	(1.71)	(1.25)	(4.96)	(3.31)	(3.63)	(0.32)
Amusic												
Mean	10 F	0 L	23.50	17.00	18.88	19.88	18.06	23.25	19.25	22.81	56.81	3.02
(SD)	6 M	16 R	(0.82)	(0.73)	(2.16)	(2.36)	(2.54)	(2.32)	(4.33)	(4.20)	(4.40)	(2.17)
***t*-Test**												
t			−1.22	−0.46	13.49	12.27	13.54	7.67	2.92	2.71	19.50	−4.77
p			0.23	0.65	< 0.001	< 0.001	< 0.001	< 0.001	0.007	0.01	< 0.001	< 0.001
Effect size												
Cohen's d			−0.43	−0.16	4.77	4.34	4.79	2.71	1.03	0.96	6.89	−1.69

F=female; M=male; L=left; R=right; age and education are in years. Scores on the six MBEA subtests (scale, contour, interval, rhythm, meter, and memory; Peretz et al., 2003) are in number of correct responses out of 30; the pitch composite score is the sum of the scale, contour, and interval scores. Pitch direction threshold (in semitones) is the mean threshold for identification of pitch direction in the speech syllable /ma/ and its complex tone analog across discrete and gliding stimuli in the same 32 participants in Liu et al. (2012b). t is the statistic of the Welch two sample t-test (two-tailed, df=30). The effect size measure Cohen's d suggests large effect sizes for the differences between amusics and controls in all MBEA subtests and pitch thresholds.