Tonal Aspects of Languages 2016
24-27 May 2016, Buffalo, New York

Impaired Vowel Discrimination in Mandarin-speaking Congenital Amusics

Xunan Huang [1, 2], *Caicai Zhang* [1, 3], *Feng Shi* [2], *Nan Yan* [1], *Lan Wang* [1]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[2] College of Chinese Language and Culture, Nankai University, China
[3] Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University, Hong Kong, China

caicai.zhang@polyu.edu.hk

Abstract

This paper investigates if individuals with amusia show deficits in the identification and discrimination of Mandarin vowels, with the aim of exploring whether the deficiency of the amusics lies in the acoustic processing of frequency, or in pitch processing. The results showed that the amusics performed comparably as the controls in vowel identification. For discrimination, both groups exhibited better discrimination for between-category pairs than within-category pairs, indicating that the amusics are not impaired in the categorical perception of vowels. However, amusics exhibited poorer accuracy than the controls in vowel discrimination across the board, irrespective of between- or within-category vowel pairs. Moreover, the participants' vowel discrimination accuracy is significantly correlated to their musical ability, as indexed by the Montreal Battery of Evaluation of Amusia (MBEA) scores. The results suggest that individuals with congenital amusia might be impaired in frequency processing in general, a deficiency broader than originally believed.

Index Terms: congenital amusia, pitch, frequency, vowel, categorical perception

1. Introduction

Enjoying music, like speaking, is an evolutionary endowment of all human beings. However, some people lack this capacity despite their efforts [1]. Those people who suffer from musical deficits are termed "congenital amusics" (amusics hereafter). Congenital amusia is usually described as a neurodevelopmental disorder that affects the processing of fine-grained musical pitch [2-5]. It is estimated that about 3-4% of individuals have this kind of impairment [6] [7]. Although the etiology for congenital amusia is not precisely known, converging evidence indicates that amusics lack the ability to detect fine-grained pitch differences. However, the systematic use of pitch is not unique to music. Pitch is also important in language. In music, pitch distinguishes different notes, whereas in languages, pitch changes intonation and a speaker's emotional status in general, and serves an additional function of determining lexical identity in tone languages.

There are cross-domain transfer effects between music and speech. Congenital amusia has been found to interfere with pitch processing in speech. For example, Tillman *et al.* [8] tested French-speaking amusics' perception of Mandarin and Thai tones and their musical analogues in a same-different paradigm, and found that the performance of amusics was inferior to that of controls for all materials. This suggested a domain-general pitch-processing deficit. On the other hand, lexical tone experience and musical pitch perception can facilitate each other. For instance, English-speaking musicians showed better performance in the identification of intact and silent-center lexical tones than English-speaking nonmusicians [9]. Moreover, comparing the performance of English-speaking nonmusicians to those of English-speaking musicians and Cantonese speakers who had minimal musical training, tone language and music training background were associated with better pitch discrimination sensitivity [10]. The cross-domain transfer might be due to the fact that music and speech commonly rely on pitch perception.

Deficits of amusia in fine-grained musical and lexical pitch processing have been extensively documented. However, the perceptual characteristics of other speech frequency components, such as formants of vowels, have not yet been examined in congenital amusia. There is an intricate relationship between pitch and vowels. Pitch depends mainly on the fundamental frequency (F0) of the sound [11], whereas vowels are primarily cued by the frequencies of the first three formants. They are both frequency-based. Additionally, it is believed that fundamental frequency (pitch) plays a major role in vowel perception. Vowels have been found to vary in intrinsic F0 such that high vowels tend to have higher intrinsic F0 than low vowels, which is closely related to vowel identification [12-13]. In order to test whether the deficit of amusics transfers to the perception of other frequency-based speech sounds, the perception of Mandarin vowels will be examined.

Additionally, amusics may also have some deficiencies in phonological processing. According to Nan *et al.* [7], Mandarin amusics were not impaired when similar tonal contrasts were carried by the same syllable, but were impaired in detecting native tonal contrasts carried by different syllables. They attributed the impairment to the difficulty of amusic individuals to filter out irrelevant variations than controls. The decomposition of phonetic segments relates to phonological processing. Wang *et al.* [14] tested the discrimination performance of three Cantonese level tones in Mandarin-speaking amusics. Those Cantonese tones were classified to familiar and unfamiliar types according to their acoustic similarity to the Mandarin tones. Results revealed that the amusics performed worse when processing speech stimuli that were native-like. This indicates that phonological inventories may affect amusics' speech discriminations.

Categorical perception is a fundamental feature of phonological processing [15], and has been extensively studied over the past 60 years. It refers to the ability of listeners to perceive continuous acoustic signals as discrete phonological categories, resulting in a sharper identification curve and better discrimination of stimuli across category boundaries than the equivalently separated stimuli within the same category [16]. Phonological labels are exclusively

 http://dx.doi.org/10.21437/TAL.2016-30

recruited in the identification task; but both auditory and phonological cues are utilized during the discrimination task [17]. In this respect, the paradigm of categorical perception also involves the ability of fine-grained auditory processing, which can be detected in the discrimination task. Specifically, formant frequencies of the equivalently separated stimuli vary in subtle ways; therefore discriminations of the stimuli contrasts require the ability to perceive tiny differences between the frequency-based formants of vowels. Categorical perception of Mandarin lexical tones in Mandarin-speaking amusics has been systematically examined by Jiang *et al.* [18]. They found that the individuals with amusia showed no improvement for between-category discrimination, indicating a lack of categorical perception of Mandarin tones.

By employing the categorical-perception paradigm, this research aims to investigate whether the impairments of pitch perception in Mandarin-speaking amusics transfers to deficits in the perception of vowels. Based on the results, we then discuss whether the deficiency lies in the lower level of acoustic processing of frequency, or in the impairment of categorical perception that relates to phonological processing. If amusics are impaired in the categorical perception of vowels, they are expected to show shallower change of identification rates across the identification boundary and pronounced impairment in the discrimination of between-category pairs. Conversely, if amusics show deficits in the lower level of frequency-based formant, they are expected to show inferior performance in vowel discrimination across the board, irrespective of between- or within-category vowel pairs.

2. Methods

2.1. Participants

Twelve Mandarin amusics (mean age=23.58 years, *SD*=2.31; average years of education=17.46 years, *SD*=1.48） and twelve matched normal controls (mean age=23.25 years, *SD*=3.60; average years of education=16.67 years, *SD*=1.91） who had not received formal musical education were recruited. The online Montreal Battery of Evaluation of Amusia (MBEA) [19] was used in the participant screening stage. Subjects who scored above 80% were classified for the control group (average score=85.17, *SD*=4.34), and those who scored below 70% were classified as amusics (average score=61.83, *SD*=5.17). Results of an independent samples *t*-test revealed that the MBEA scores were significantly different between the two groups ($p < .001$).

2.2. Materials

The stimuli were synthesized from natural productions of /ɤ 55/ (婀 'fair') and /u 55/ (乌 'black') by a native male speaker from Beijing. Based on the recorded utterance of /ɤ/, nine speech stimuli were synthesized by simultaneously varying F1 and F2 with equidistant values using Praat [20]. All nine speech stimuli are 450 ms long. The formant frequency manipulation is illustrated in Fig. 1. Stimulus 1 represents typical /ɤ 55/ and stimulus 9 represents typical /u55/.

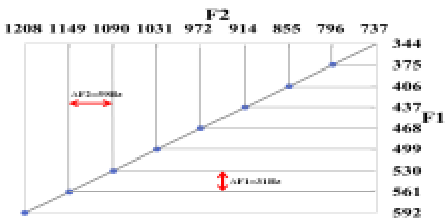

Fig. 1: *The schematic diagram of stimulus continuum: the stimuli are synthesized from /ɤ/ to /u/ with ΔF1≈31Hz and ΔF2≈59Hz in every step.*

2.3. Procedure

All participants took part in three tasks: vowel identification task, vowel discrimination with two-step difference (DC task1 hereafter) and a more elaborate discrimination task with one-step difference (DC task2 hereafter). Subjects were instructed to press 1 to represent /ɤ/ and press 2 to represent /u/ in the identification task, and were instructed to respond by pressing button 1 to indicate same vowel pairs and button 2 to indicate different vowel pairs in the discrimination task as quickly and accurately as possible. The inter-stimulus-interval was set to 500 ms, and the maximum reaction time was 2000 ms. Stimuli were presented in random order and repeated 9 times within each task. Each task was divided into three blocks with a 20-second break in between. There were 27 trials in every vowel identification block, 72 trials in DC task1, and 75 trials in DC task2.

2.4. Data Analysis

Given a particular stimulus, the identification score was defined as the percentage of responses with which participants identified that stimulus as being either /ɤ/ or /u/. In the following identification curves, only the percentage of /u/ response is presented, and the percentage of /ɤ/ is 100% minus that of /u/. Boundary position and boundary width were assessed by Profit analyses of individual identification curves [21]. The boundary position was defined as the 50% crossover point, while the boundary width was defined as the linear distance between the 25th and 75th percentiles [22]. Discrimination accuracy of each pair was calculated by hits (number of correct responses for different-token trials / number of different-token trials) minus false alarms (F.A.; number of incorrect responses for same-token trials / number of same-token trials), following the previous study [18].

3. Results

3.1. Vowel Identification

Identification curves of /u/ responses for the two groups are shown in Fig. 2. A two-way repeated measures ANOVA was conducted with group (control and amusic) as the between-subject factor, and stimulus step (9 steps) as the within-subject variable. There was neither a significant main effect of groups [F (1, 22) = 1.268, p=.272], nor an interaction effect between group and response rate [F (8, 176) = 1.263, p=.266], but there was a significant main effect of stimulus number [F (8, 176) = 875.394, p< .001].

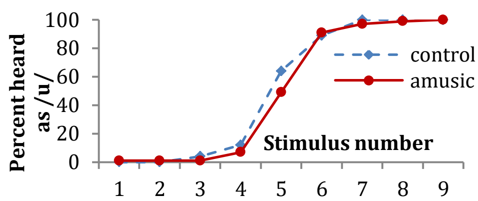

Fig. 2: *The identification curves are derived from the average percentage of /u/ in every stimulus, and / ɤ / makes up the other part of a hundred percentages.*

Table 1. *Average categorical boundary position and boundary width for each group with their standard deviations in the brackets.*

	Boundary position	Boundary width
Control	4. 79 (0.54)	0. 88 (0.32)
Amusic	5.04 (0.55)	0. 96 (0.58)

The estimated average boundary position and boundary width are shown in Table 1. Independent-samples *t*-tests were conducted to examine whether perceptual boundary positions or boundary widths were significantly different between two groups. Neither boundary position (t=-1.101, p=.284) nor boundary width (t=-.367, p=.718) was significantly different between amusics and controls.

Furthermore, correlation between individual identification boundary width and the scores of MBEA was examined with the Spearman correlation test, and the results revealed no significant correlation [r (24) =.24, p=.910]. Correlations between the boundary position and the scores of MBEA were also not significant [r (24) = - .177, p=.407].

3.2. Vowel Discrimination

Fig. 3 displays the mean scores of hits minus F.A. of each stimulus pair in DC task1 (Fig. 3a) and DC task2 (Fig. 3b). Two-way repeated measures ANOVAs were conducted on discrimination accuracies of DC task1 and DC task2 respectively, with stimulus pairs as the within-subject variable, and group (control and amusic) as the between-subject factor.

Significant main effects of group were found in both DC task1 [F (1, 22) = 5.974, p < .05] and DC task2 [F (1, 22) = 7.322, p < .05], which indicated that the overall discrimination accuracies were significantly different in the two groups, with the discrimination accuracies of amusics being lower than those of the controls, as shown in Fig. 3. The main effects of stimulus pairs were also significant in both DC task1 [F (6, 132) = 17.338, p < .001] and DC task2 [F (7, 104) = 14.531, p < .001], but no interaction effects of stimulus pairs with groups were found in either of the tasks.

Stimulus pairs were then divided into between-category and within-category comparisons based on the position of the identification boundary for each subject. For example, if the participant' s classification boundary was at 4.5, then the scores for stimulus pairs 3–5 and 4–6 would be averaged and coded as between-category comparisons, while the remaining comparisons would be coded as within-category [18]. Fig. 4 shows the individual and mean scores of hits-F.A. for both groups on between- and within- category discriminations of DC task1 (a) and DC task2 (b). A three-way repeated measures ANOVA was conducted with group (control and amusic) as the between-subject factor, and difference size (two-step pairs and one-step pairs) and category status (between-category and within-category) as the within subject variables.

There was a significant main effect of group [F (1, 22) = 7.992, p < .05] with discrimination accuracy of controls being significantly higher than that of amusics, and a significant main effect of category [F (1, 22) = 64.431, p < .001] with discrimination accuracy between-category being significantly higher than that of within-category. The significant main effects of group and category indicated that although the discrimination accuracies of amusics are lower, they showed the enhancement in performance when vowel pairs crossed the boundary positions. Additionally, there was a significant main effect of size difference [F (1, 22) = 62.188, p < .001], with discrimination accuracy of two-step pairs being significantly better that of one-step pairs. But there were no significant interaction effects between group and other factors.

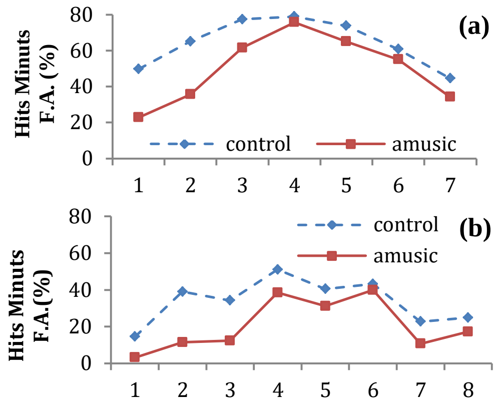

Fig. 3： *The abscissas represent stimulus pairs: the scores of stimulus pair 1 in (a) and (b) are measured by average discrimination accuracies of 1-3 and 3-1; 1-2 and 2-1respectively.*

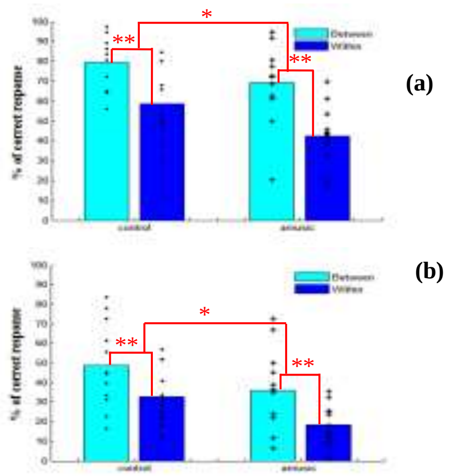

Fig. 4： *Individual and mean discrimination scores in DC1 (a) and DC2 (b) for between- category and within-category pairs; average and individual scores are represented by columns and dots respectively.*

In addition, correlations between the participants' MBEA scores and their discrimination accuracies in two tasks were conducted. Significant correlations were found in both tasks. For task1, MBEA scores were positively correlated with the discrimination accuracies [r (24) =.502, p<.05]. Similarly, the correlations of MBEA scores with that in DC task 2 were also significant [r (24) =.571, p<.005].

4. Discussion

As mentioned in the introductory part, vowel and pitch are both frequency-based. Conjunction evidences of this research on impaired formant-processing and previous studies in pitch-processing deficiencies of amusics [8] may indicate that individuals with amusia are impaired in frequency processing in general. This is broader than the deficiency originally believed.

In the present study, data on the identification tasks do not show any significant differences between the controls and the amusics, as shown in Fig. 2 and Table 1. Moreover, discrimination accuracies of between-category pairs are significantly higher than those of within-category pairs for both amusics and controls in the discrimination tasks. Similar identification performances and cross-boundary peak for both controls and amusics in this research manifest that categorical perception of vowels are not impaired in amusics. However, the discrimination accuracy of amusics is significantly lower than that of controls across the board in this experiment. This is in line with the latter hypothesis that the amusics show deficits in the lower level of frequency-based formant processing, rather than the impairment in the categorical perception of Mandarin vowels.

It is important to note that amusics do not show impaired categorical perception of Mandarin vowels. This is in contrast to what was observed on lexical tone perception by Jiang *et al.* [18]. In the categorical perception of Mandarin tones, amusics showed no improvement in discriminating pairs that crossed the boundary; therefore they estimated that amusics exhibited deficiencies in the categorical perception of Mandarin tones. It could mean that lexical tone impairment is more severe, impairing linguistic categorical processing. But for vowels, though there are some deficits, it's more lower-level, confined to acoustic processing.

We argue that the richness of formant cues facilitates vowel perception. Amusics can detect speech differences better, such as intonations [24], when there are other cues to aid the perception. In the perception of lexical tones, only F0 was varied, whereas in the current study, the frequencies of F1 and F2 were manipulated simultaneously. Compared with those tone stimuli, vowel stimuli of this research have richer cues. This may provide more information that facilitates the categorical perception of vowels. In order to further explore whether deficiencies of pitch and vowel perception lay in different levels, future within-subjects experiments with both pitch and vowel stimulus continua will be conducted.

5. Conclusions

By comparing the categorical perception that involves both identification and discrimination between controls and amusics, we found impaired performance in vowel discrimination but normal performance in vowel identification in the Mandarin amusics. For discrimination, amusics showed inferior performance in discrimination of vowels. However, the amusics are not selectively impaired in the discrimination of between-category vowel pairs, and exhibited poorer accuracy in vowel discrimination across the board irrespective of between- or within-category vowel pairs. This suggests that individuals with congenital amusia might be impaired in frequency processing in general.

6. Acknowledgements

This work was supported by grants from the National Outstanding Youth Science Fund Project: "1150040716 Production and perception of lexical tone and music in Cantonese-speaking amusics".

7. References

[1] Peretz, I. and Hyde, K. L., "What is specific to music processing? Insights from congenital amusia", Trends in Cognitive Sciences, 7(8):362-367, 2003.

[2] Peretz, I., "Brain specialization for music. New evidence from congenital amusia", Annals of the New York Academy of Sciences, 930(4):153-165, 2001.

[3] Peretz, I., Ayotte, J., Zatorre, R.J. and Mehler, J. *et al.*, "Congenital amusia: a disorder of fine-grained pitch discrimination", Neuron, 33:185, 2002.

[4] Hyde, K. L. and Peretz, I., "Brains that are out of tune but in time", Psychological Science, 15(5):356-360, 2004.

[5] Jiang, C. M., and Yang, Y., "Fine-grained pitch discrimination in congenital amusics with Mandarin Chinese", Music Perception, 28(5):519-526, 2011.

[6] Kalmus, H. and Fry, D. B., "On tune deafness (dysmelodia): frequency, development, genetics and musical background", Annals of Human Genetics, 43(4): 369–382, 1980.

[7] Nan, Y., Sun, Y. and Peretz, I., "Congenital amusia in speakers of a tone language: association with lexical tone agnosia", Brain, 133(6):2635-2642, 2010.

[8] Tillmann, B., Burnham, D., and Nguyen, S. *et al.* "Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages", Auditory Cognitive Neuroscience, 2(4): 120, 2011.

[9] Lee, C. Y., Hung, T. H., "Identification of Mandarin tones by English-speaking musicians and nonmusicians", Journal of the Acoustical Society of America 124(5): 3235-3248, 2008.

[10] Bidelman, G. M., Stefanie, H. and Sylvain, M., "Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for directionality between the domains of language and music", Plos One, 8(4):e60676, 2013.

[11] Hartmann, W. M., and Bartram, J. F. "Signals, Sound, and Sensation. Signals, sound, and sensation", American Institute of Physics, 579, 1997.

[12] William, S-Y. W., "The Many Uses of F0", A Revised Version at the Kyoto Speech Communication Symposium, 1968.

[13] Turner, P. and Verhoeven, J., "Intrinsic vowel pitch: a gradient feature of vowel systems", The 17th International Congress of Phonetic Sciences, 2011.

[14] Wang, X., and Peng, G., "Phonological processing in Mandarin speakers with congenital amusia", Journal of the Acoustical Society of America, 136: 3360-3370, 2014.

[15] Harnad, S. R. "Categorical Perception: The Groundwork of Cognition (eds.)", New York, NY: Cambridge University Press, 1987.

[16] Repp, B.H., "Categorical perception: Issues, methods, findings", In: N.J. Lass (Ed.), Speech and language: Advances in basic research and practice (Vol. 10) (pp. 243–335). New York: Academic Press, 1984.

[17] Xu, Y. S., Gandour J. T., Francis, A. L. "Effects of language experience and stimulus complexity on the categorical perception of pitch direction", Journal of the Acoustical Society of America, 120(2):1063-74, 2006.

[18] Jiang, C. M., Hamm, J. P., and Lim, V. K. *et al.* "Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics", Memory & Cognition, 40(7):1109-1121, 2012.

[19] Peretz I., "On-Line identification of congenital amusia". Music Perception, 25(4):331-343, 2008.

[20] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer". (Computer program). Version 5.3.56, retrieved 15 Sep 2013 from http://www.praat.org/, 2013.

[21] Finney, D. J., "Probit Analysis" (3rd). Cambridge: Cambridge University Press, 1971.

[22] Peng, G., Zheng, H. Y., Gong, T., *et al.*, "The influence of language experience on categorical perception of pitch contours", Journal of Phonetics, 38: 616-624, 2010.

[23] Wu, X., and Lin, H., "Perception of Mandarin tones by Mandarin and English listeners", Journal of Chinese Language and Computing, 18:175–187, 2008.

[24] Patel, A. D., "Speech Intonation perception deficits in musical tone deafness (congenital amusia)", Music Perception, 25(4): 357-368, 2008.