INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Effect of noise on lexical tone perception in Cantonese-speaking amusics

Jing Shao [1], *Caicai Zhang* [1,2], *Gang Peng* [1,2,3], *Yike Yang* [1], *William S-Y. Wang* [1,2]

[1] Department of Chinese and Bilingual Studies, the Hong Kong Polytechnic University
[2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[3] Department of Linguistics and Modern Languages, the Chinese University of Hong Kong
`jingshao@polyu.edu.hk, caicai.zhang@polyu.edu.hk, gpeng@cuhk.edu.hk,`
`yike.yang@polyu.edu.hk, wsywang@polyu.edu.hk`

Abstract

Congenital amusia is a neurogenetic disorder affecting musical pitch processing. It also affects lexical tone perception. It is well documented that noisy conditions impact speech perception in second language learners and cochlear implant users. However, it is yet unclear whether and how noise affects lexical tone perception in the amusics. This paper examined the effect of multi-talker babble noise [1] on lexical tone identification and discrimination in 14 Cantonese-speaking amusics and 14 controls at three levels of signal-to-noise ratio (SNR). Results reveal that the amusics were less accurate in the *identification* of tones compared to controls in all SNR conditions. They also showed degraded performance in the *discrimination,* but less severe than in the identification. These results confirmed that amusia influences lexical tone processing. But the amusics were not influenced more by noise than the controls in either identification or discrimination. This indicates that the deficits of amusia may not be due to the lack of native-like language processing mechanisms or are mechanical in nature, as in the case of second language learners and cochlear implant users. Instead, the amusics may be impaired in the linguistic processing of native tones, showing impaired tone perception already under the clear condition.

Index Terms: congenital amusia, Cantonese, lexical tones, SNR, discrimination, identification.

1. Introduction

Congenital amusia (amusia, hereafter) is a lifelong deficit in the processing of musical pitch in the absence of brain injury [2], [3], [4]. It has been estimated that amusia affects around 3-4% of the population for speakers of both tonal and non-tonal languages [5]. Individuals with amusia have difficulty in making fine-grained pitch discriminations [6], and in processing pitch-change direction [7]. They also demonstrated impaired memory for pitch [8]. Evidence has shown that amusia is not a music-specific disorder, but also transfers to the language domain and influences pitch processing in speech, including linguistic and emotional prosody processing, and speech intonation imitation [9], [10]. The deficits in amusia were also found affecting lexical tone processing. Nguyen et al. [11] found that French-speaking amusics performed significantly worse than the control group on the discrimination of Mandarin lexical tones, indicating a transfer of musical deficits to lexical tone perception. In another study [12], French-speaking amusics also showed impairment in the processing of Mandarin and Thai lexical tones. These results indicated that the discrimination of lexical tones was problematic in non-tonal language speakers with amusia.

In tonal languages, tone changes systematically differentiate word meanings. It is hypothesized that tonal language experience might compensate for musical pitch deficits, such that amusical speakers of a tonal language may show intact sensitivity to pitch changes that occur in their native language [13]. Whereas one study shows that the prevalence rate of amusia might be lower in the Cantonese-speaking population than in non-tonal language speakers [14], providing some support for this hypothesis, many studies suggested that this hypothesis might not be true. For example, Nan et al. [4] found that nearly half of 22 Mandarin-speaking amusics showed impairments in both identification and discrimination of Mandarin lexical tones. Jiang et al. [15] found that Mandarin-speaking amusics failed to show enhancement in accuracy for discriminating tone pairs that crossed the classification boundary, suggesting a lack of categorical perception. Wang and Peng [16] found that Mandarin-speaking controls discriminated non-native Cantonese level tones more accurately when the tones were carried by native Mandarin syllables than by Cantonese syllables. However, the amusics failed to benefit from such facilitating effect. It suggests that language familiarity influences the discrimination performance in amusics.

The aforementioned studies suggest that tonal language speakers with amusia also show impairment in lexical tone perception. However, little is known about the effect of *noise* on lexical tone perception by tonal language speakers with amusia. It has been documented that unfavorable listening conditions impact speech perception in important ways. For example, second language (L2) learners are particularly prone to the influence of noise, performing less accurately in speech perception under noisy listening conditions, despite that they performed similarly well with monolinguals under quiet conditions [17]. Noise condition also affects the cochlear implant (CI) users, who struggled to understand speech in noise in spite of excellent speech perception in quiet [18]. This leads to the question of whether amusics are vulnerable to the influence of noise, like L2 learners and CI users. To this end, we aim to investigate the effect of noise on lexical tone perception in Cantonese-speaking amusics. We tested Cantonese-speaking amusics and controls' tone perception in identification and discrimination tasks with Cantonese tones embedded in noise.

http://dx.doi.org/10.21437/Interspeech.2016-891

2. Method

2.1. Participants

Fourteen amusics and 14 controls participated in the experiment. All participants were native speakers of Hong Kong Cantonese, right-handed, with no hearing impairment, and no reported history of musical training. Amusics and controls were selected according to the Online Identification Test of Congenital Amusia [19]. It consists of three parts: out-of-key, offbeat and mistuned tests. All amusics scored 70 or lower, and all controls scored 80 or higher. Characteristics of amusics and controls are summarized in Table 1.

Table 1. *Demographic characteristics of amusic and control participants in the behavioral experiment.*

	Amusics	**Controls**
Male / Female (Total)	5/ 9 (14)	5 / 9 (14)
Age (range)	21.3 ± 2.6 yr (18.8-28.6 yr)	21.9 ± 2.3yr (18.8-27.3 yr)
Test of Congenital Amusia		
Out-of-key (SD)	65.5 (7.1)	89.4 (5.1)
Offbeat (SD)	71.9 (9.8)	87.5 (7.4)
Mistuned (SD)	59.2 (5.9)	92.3 (7.9)
Global score (SD)	65.4 (3.9)	89.6 (4.1)

2.2. Stimuli

The stimuli were 12 words contrasting six Cantonese unchecked tones (high level tone (T1)-/55/, high rising tone (T2)-/25/, mid level tone (T3)-/33/, extra low level/low falling tone (T4)-/21/, low rising tone (T5)-/23/, low level tone (T6)-/22/) on two base syllables (/ji/ and /fu/). A female Cantonese speaker was recorded reading aloud the words in isolation for 10 times. For each word, one clear token was selected, and normalized in duration to 500 ms and in mean intensity to 70 dB using Praat [20]. The F0 trajectroy is preserved after the normalization. The same set of words were then embedded in multi-talker babble noise at three levels of signal-to-noise ratio (SNR): clear (no noise), moderate noise (SNR 0 dB) and severe noise (SNR −10 dB).

2.3. Procedure

The experiment included an identification task and a discrimination task, generated by E-prime 2.0. The identification task always preceded the discrimination task. For each task, there were three blocks corresponding to the three noise levels, and the presentation order was counter balanced across the participants. Before each task, subjects were given a few practice trials (/a/ carrying six tones) to familiarize them with the procedures.

In the identification task, each set of six words (/ji/ and /fu/) was presented in a sub-block, to avoid confusion. Six words in a set were repeated twice and presented randomly in a sub-block. In each trial, a fixation first occurred on the computer screen for 500 ms, followed by the presentation of a 500-ms stimulus via headphones. The subjects were instructed to identify the tone of the word by pressing buttons 1-6 on a keyboard within 5 seconds. Subjects were given a list of words contrasting the six tones beforehand to familiarize them with the tone distinction and facilitate tone identification. Half of the participants identified /ji/ stimuli first and the other half identified /fu/ stimuli first.

In the discrimination task, six words in each set (/ji/ and /fu/) were grouped into 15 different tone pairs and 6 same tone pairs. Each set was presented in a sub-block, to avoid confusion. Within a sub-block, different tone pairs were repeated twice and same tone pairs were repeated five times, generating equal number of different and same tone pairs, which were intermixed and randomly presented. In each trial, a fixation first occurred on the computer screen for 500 ms, followed by the presentation of two 500-ms stimuli separated by an inter-stimulus-interval of 500 ms via the headphones. The subjects were instructed to judge whether the two words carried the same tone or different tones by pressing "left arrow" (same) and "right arrow" (different) on a keyboard within 3 seconds. Half of the subjects listened to the /ji/ stimuli first and the other half listened to the /fu/ stimuli first. Accuracy and response time (RT) were collected.

2.4. Data analysis

For the identification, accuracy and reaction time (RT) were analyzed. Accuracy was the percentage of correctly identified trials for each noise condition per subject. Arcsine transformation was then applied to the percentage data. As for RT analysis, incorrect trials were disregarded, as were trials exceeding three SDs of each condition (0.2%). For the discrimination results, the sensitivity index d' and RT were analyzed. The d' [21] was computed as the z-score value of the hit rate ("different" responses to different tone pairs) minus that of the false alarm rate ("different" responses to same tone pairs) for each noise condition per subject. As for RT analysis, incorrect trials were disregarded, as were trials exceeding three SDs of each condition (1.8%).

Group × *noise level* × *tone* repeated measures ANOVAs were conducted on the identification accuracy, identification RT, discrimination sensitivity d' and discrimination RT, using Statistical Package for the Social Sciences (SPSS) [22].

3. Results

Figure 1 and 2 show the identification accuracy and d' scores under three SNR conditions. For the identification accuracy, there were significant main effects of *group* ($F(1, 26) = 19.050, p < 0.001$), *noise level* ($F(1.427, 37.098) = 38.643, p < 0.001$), and *tone* ($F(4.234, 110.093) = 21.878, p < 0.001$), as well as significant two-way interactions between *group* and *noise level* ($F(2, 52) = 5.751, p = 0.006$), and between *tone* and *noise level* ($F(10, 260) = 2.575, p = 0.005$). First, we analyzed the *group* by *noise level* interaction. One-way ANOVAs were conducted to examine the effect of *noise* within each group. For the amusics, the effect of *noise* was significant ($F(2, 249) = 7.251, p = 0.001$). Tukey-corrected post hoc tests showed that the accuracy in the severe condition was significantly lower than the clear condition ($p = 0.002$) and the moderate condition ($p = 0.008$). No other effects were significant. For the controls, there was also a significant effect of noise ($F(2, 249) = 35.733, p < 0.001$). Post hoc tests showed that the accuracy in the severe condition was significantly lower than the clear condition ($p < 0.001$) and the moderate condition ($p < 0.001$). Independent samples t-tests were conducted to examine the effect of *group* within each noise level. There were significant group differences in the clear condition ($t(166) = 5.391, p < 0.001$), the moderate condition ($t(166) = 6.087, p < 0.001$), and the severe condition ($t(166) = 2.131, p$

$= 0.035$), where the controls always exhibited higher accuracy than amusics. Although the controls always performed more accurately, the difference between controls and amusics shrunk from clear to severe conditions. Second, we analyzed the *tone* and *noise level* interaction. One-way ANOVAs were conducted to examine the effect of *noise* within each tone. Main effects of *noise* were found on T1 ($F(2, 81) = 32.188$, $p < 0.001$), T2 ($F(2, 81) = 3.588$, $p = 0.032$), T4 ($F(2, 81) = 5.405$, $p = 0.006$), T5 ($F(2, 81) = 6.254$, $p = 0.003$) and T6 ($F(2, 93) = 5.929$, $p = 0.004$). Post hoc results showed that for T1, the identification accuracy in the severe condition was significantly lower than the clear condition ($p < 0.001$) and the moderate condition ($p < 0.001$), similar results were also found for T4 ($p = 0.01$; $p = 0.026$) ; for T2 and T6, the accuracy in the severe condition was significantly lower than the moderate condition ($p = 0.026$; $p = 0.003$); for T5, the accuracy in the severe condition was significantly lower than the clear condition ($p = 0.002$). One-way ANOVAs were conducted to investigate the effect of *tone* within each noise level. Significant effects were obtained for the clear condition ($F(5, 162) = 7.670$, $p < 0.001$), the moderate condition ($F(5, 162) = 8.719$, $p < 0.001$) and the severe condition ($F(5, 162) = 3.714$, $p = 0.003$). Post hoc results revealed that under the clear condition, the identification accuracy for T1 was significantly higher than T3 ($p = 0.001$), T5 ($p < 0.001$) and T6 ($p < 0.001$), while the accuracy for T4 was significantly higher than T5 ($p = 0.022$). Under the moderate noise condition, the accuracy of T1 was significantly higher than T2 ($p = 0.001$), T3 ($p < 0.001$), T4 ($p = 0.039$), T5 ($p < 0.001$) and T6 ($p < 0.001$), and T4 also received significantly higher accuracy than T6 ($p = 0.030$). Under the severe condition, T4 was recognized significantly better than T6 ($p = 0.039$). It can be observed that in the clear and moderate noise conditions, there is a tendency that T1 is more accurately identified compared to other level tones, and that T4 is relatively easier to identify than T6. Moreover, both being rising tones, T2 is more accurately identified than T5.

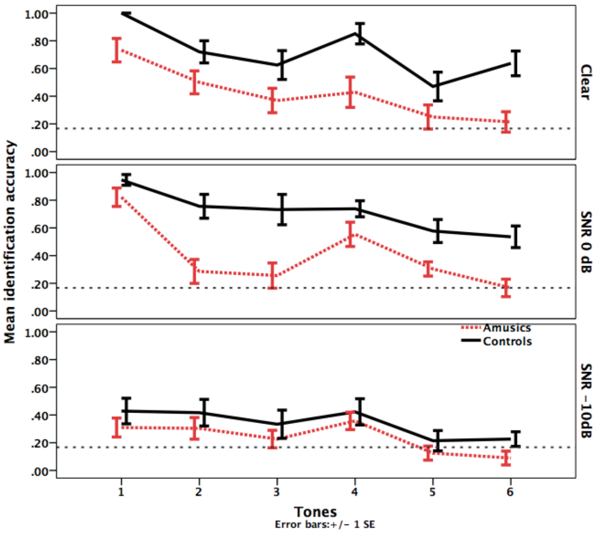

Figure 1: *The untransformed identification accuracy. The grey dotted line indicates chance-level accuracy (0.167).*

For the identification RT, no effects were significant.

For the d', there was a significant main effect of *group* ($F(1, 26) = 8.954$, $p = 0.006$), where the d' of the amusics was significantly lower than that of controls. There were also significant main effects of *noise level* ($F(2, 52) = 354.864$, $p < 0.001$), and *tone pairs* ($F(14, 364) = 19.774$, $p < 0.001$), as well as significant interaction between *noise level* and *tone pairs* ($F(11.217, 291.651) = 1.357$, $p = 0.004$). First, one-way ANOVAs were conducted to examine the effect of *tone pairs* within each noise level. Significant effects were obtained for the clear condition ($F(14, 405) = 8.469$, $p < 0.001$) and the moderate condition ($F(14, 405) = 9.328$, $p < 0.001$). Post hoc results revealed that under the clear condition, the difference mainly existed between the pair T2/T5 and the other tone pairs, with the d' score for the T2/T5 pair being significantly lower than others ($ps < 0.001$). When the noise was moderate, the T2/T5 pair and T3/T6 pair received significantly lower d' score than all other pairs ($ps < 0.001$; $ps < 0.037$), whereas the difference between T2/T5 and T3/T6 was not significant ($p = 0.66$). These results suggested that the T2/T5 and T3/T6 pairs are of particular difficulties for the Cantonese speakers, and the difficulty was most pronounced in clear and moderate noise conditions. Second, One-way ANOVAs were conducted to examine the effect of *noise level* within each tone pair. Significant effects were found for all the 15 pairs ($ps < 0.001$). Post hoc results indicated that the d' score under the severe condition was always significantly lower than the clear and moderate conditions ($p < 0.001$), with one exception of the T3/T6 pair, as the score in the clear condition was also higher than in the moderate condition ($p = 0.035$) for this pair.

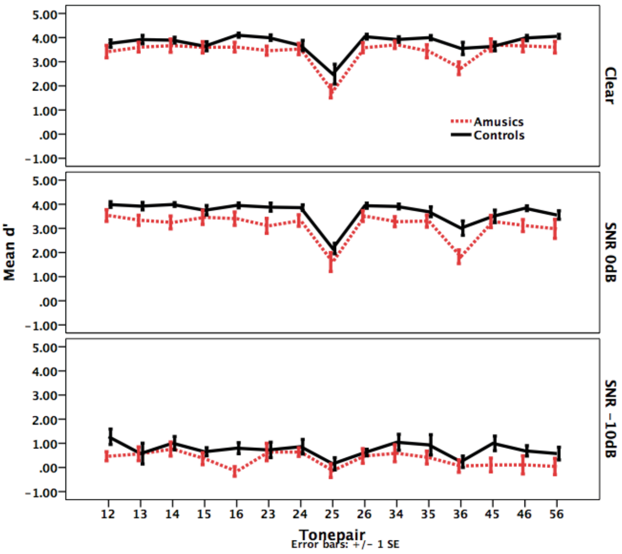

Figure 2: *Discrimination sensitivity index d'.*

As for the discrimination RT, there was a significant main effect of *group* ($F(1, 26) = 6.369$, $p = 0.018$), where amusics responded significantly more slowly than the controls in the discrimination task, indicating that they are less sensitive to the tone differences than the normal controls. There was also a significant two-way interaction between *tone pairs* and *SNR levels* ($F(40, 1040) = 2.000$, $p < 0.001$). No other effects were significant. One-way ANOVA was conducted to examine the effect of *noise* within each tone pair. Significant effects were found in T2/T4 pair ($F(2, 80) = 3.415$, $p = 0.038$), T2/T6 pair ($F(2, 79) = 3.715$, $p = 0.029$), T3/T4 pair ($F(2, 79) = 3.690$, $p = 0.029$), T4/T4 pair ($F(2, 81) = 5.649$, $p = 0.005$), T4/T6 pair ($F(2, 79) = 5.663$, $p = 0.005$), and T6/T6 pair ($F(2, 81) = 3.419$, $p = 0.038$). Post hoc tests revealed that for tone pairs

T2/T4, T2/T6, T3/T4 and T6/T6, the RT in the severe condition was significantly slower than the clear condition, while for pairs T4/T4 and T4/T6, the RT in the severe condition was significantly slower than the clear and moderate condition.

4. Discussion

The present study examined the effect of noise on lexical tone identification and discrimination by Cantonese-speaking amusics and controls. Overall, our results showed that Cantonese-speaking amusics were impaired in lexical tone perception, confirming that the pitch-processing deficit in amusia is not limited to music, but also transfers to the language domain. Tonal language experience did not compensate for the deficit associated with amusia.

Our results revealed that Cantonese-speaking amusics were impaired in both identification and discrimination of the lexical tones. This finding echoes with the results found in Mandarin-speaking amusics [4]. The amusics identified the tones with significantly lower accuracy compared to the controls, and their discrimination sensitivity was also decreased compared with the controls. Moreover, the amusics took a significantly longer time to make correct responses than the controls. As response time is also an important index of perceptual sensitivity, it can be concluded that the amusics had to make more cognitive effort to perform the discrimination task, and thus showed less perceptual sensitivity to tone differences than the controls.

It is worth noting that although the amusics performed less accurately in tone categorization than the controls in all three SNR levels, the group difference tends to reduce from the clear and the moderate conditions to the severe condition. This indicates that the amusics are *not* more vulnerable to the influence of noise than the controls. This finding is consistent with a previous study [23], in which Mandarin-speaking amusics were tested on the intelligibility of Mandarin sentences with natural and flat fundamental frequency contours under four SNR conditions (no noise, SNR +5, 0, -5 dB). The amusics showed more degraded comprehension performance than the controls across the board. Both normal and amusics were largely unaffected by flattened pitch contour in quiet and extremely noisy conditions (-5 dB). But in moderately noisy conditions, both groups exhibited significant loss of speech intelligibility for sentences with flattened relative to natural pitch contour. These results suggested that certain noise conditions influence sentence intelligibility, but noise condition did not impact the amusics more than the normal controls in sentence processing. Our results provided extra evidence that noise does not affect the amusics more than the controls in lexical tone perception.

The fact that the amusics were not further impaired in unfavorable listening conditions is different from the findings in L2 speech perception and CI users. This discrepancy might reflect different kinds of deficits in the three groups of listeners. L2 learners are more vulnerable to the influence of noise, as they have not established the optimal language representations and processing mechanisms like native speakers to reliably differentiate the L2 contrasts, and this deficit is more pronounced in the noisy conditions. The CI users are impaired mechanically in hearing, and the aid of the CI is likely weakened by noise. In contrast, the deficits in amusia are not due to the lack of native-like language processing mechanisms or are mechanical in nature. Instead, the amusics may be impaired in the linguistic representations of native tones, which leads to degraded tone identification performance even under the clear condition, and this deficit does not appear to be reinforced by unfavorable listening conditions.

It is notable that the perception of the six tones was not impaired equally. Among them, T1 is of perceptual saliency and was identified with the highest accuracy regardless of the SNR conditions, which echoed previous findings [24]. T4 was also more or less accurately identified. T1 and T4 received higher accuracy, possibly because they mark the periphery of the tone space. In contrast, T3 and T6, T2 and T5 were less accurately recognized. Corresponding to these identification patterns, discrimination results revealed that the pairs T2/T5 and T3/T6 were extremely difficult to discriminate for both groups. It is not surprising that these two pairs are most difficult, because they are acoustically highly similar (high rising vs. low rising; mid level vs. low level), machines also made more errors on these two pairs when recognizing tones in continuously spoken Cantonese [25].

Interestingly, T2/T5 and T3/T6 were reported undergoing merging in Hong Kong Cantonese [26], [27]. The phenomenon of tone merger was first reported in 1974 [28], which shows that T3 and T6, and T2 and T5 were readily confused with each other perceptually. Confusion in production for these two pairs has also been reported [29]. Recent studies provided further evidence on the tone merger [24], [30], [31], showing that the T2/T5 pair and the T3/T6 pair are the main source of tone merger in Hong Kong Cantonese. It should be noted that the tone merger is unfolding in different word pairs in different speakers, affecting some individuals but not others. The confusion patterns shown by the amusics in the current study resemble the perceptual patterns of individuals who exhibit tone merger. This leads to the question of a possible link between amusia and tone merger, which should be investigated in future research.

5. Conclusions

The results of the current study revealed that Cantonese speakers with amusia showed degraded performance in tone identification and discrimination. Tones that of similar acoustic characteristics were of particular difficulties for the amusics. However, the deficit in pitch processing in amusia does not further reinforced in adverse listening conditions. The amusics' confusions on T2/T5 and T3/T6 also provide some initial evidence on the link between amusia and tone merger in Hong Kong Cantonese.

6. Acknowledgements

This study was supported by the PolyU Start-up Fund for New Recruits, and the Research Grant Council of Hong Kong (GRF: 14408914). We thank Ms. Xiao Wang, and Mr. Steve Ka-Hong Wong for their help with data collection.

7. References

[1] S. Jin and C. Liu, "English sentence recognition in speech-shaped noise and multi-talker babble for English-, Chinese-, and Korean-native listeners", *The Journal of the Acoustical Society of America*, vol. 132, no. 5, p. EL391, 2012.

[2] I. Peretz "Brain Specialization for Music", *Annals of the New York Academy of Sciences*, vol. 930, no. 1, pp. 153-165, 2006.

[3] J. Ayotte, I. Peretz and K. Hyde, "Congenital amusia: A group study of adults afflicted with a music-specific disorder", *Brain*, vol. 125, no. 2, pp. 238-251, 2002.

[4] I. Peretz, J. Ayotte, J. Zatorre, J. Mehler, P. Ahad, B. Penhune and B. Jutras, "Congenital Amusia: A Disorder of Fine-Grained Pitch Discrimination". *Neuron,* vol. 33, no. 2, pp.185–191. 2002.

[5] Y. Nan, Y. Sun and I. Peretz, "Congenital amusia in speakers of a tone language: association with lexical tone agnosia", *Brain*, vol. 133, no. 9, pp. 2635-2642, 2010.

[6] J. Foxton, "Characterization of deficits in pitch perception underlying 'tone deafness'", *Brain*, vol. 127, no. 4, pp. 801-810, 2004.

[7] F. Liu, A. Patel, A. Fourcin and L. Stewart, "Intonation processing in congenital amusia: discrimination, identification and imitation", *Brain*, vol. 133, no. 6, pp. 1682-1693, 2010.

[8] B. Tillmann, K. Schulze and J. Foxton, "Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds", *Brain and Cognition*, vol. 71, no. 3, pp. 259-264, 2009.

[9] A. Patel, M. Wong, J. Foxton, A. Lochy and I. Peretz, "Speech intonation perception deficits in musical tone deafness (congenital amusia)", *Music Perception: An Interdisciplinary Journal*, vol. 25, no. 4, pp. 357-368, 2008.

[10] C. Jiang, J. Hamm, V. Lim, I. Kirk and Y. Yang, "Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese", *Neuropsychologia*, vol. 48, no. 9, pp. 2630-2639, 2010.

[11] S. Nguyen, B. Tillmann, N. Gosselin and I. Peretz, "Tonal Language Processing in Congenital Amusia", *Annals of the New York Academy of Sciences*, vol. 1169, no. 1, pp. 490-493, 2009.

[12] B. Tillmann, D. Burnham, S. Nguyen, N. Grimault, N. Gosselin and I. Peretz, "Congenital Amusia (or Tone-Deafness) Interferes with Pitch Processing in Tone Languages", *Frontiers in Psychology*, vol. 2, 2011.

[13] L. Stewart, "Congenital amusia", *Current Biology*, vol. 16, no. 21, pp. R904-R906, 2006.

[14] P. Wong, V. Ciocca, A. Chan, L. Ha, L. Tan and I. Peretz, "Effects of Culture on Musical Pitch Perception", *PLoS ONE*, vol. 7, no. 4, p. e33424, 2012.

[15] C. Jiang, J. Hamm, V. Lim, I. Kirk and Y. Yang, "Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics", *Mem Cogn*, vol. 40, no. 7, pp. 1109-1121, 2012.

[16] X. Wang and G. Peng, "Phonological processing in Mandarin speakers with congenital amusia", *The Journal of the Acoustical Society of America*, vol. 136, no. 6, pp. 3360-3370, 2014.

[17] D. Tabri, K. Chacra and T. Pring, "Speech perception in noise by monolingual, bilingual and trilingual listeners", *Int J Lang Commun Disord*, p. 101005020050084, 2010.

[18] A. Srinivasan, M. Padilla, R. Shannon and D. Landsberger, "Improving speech perception in noise with current focusing in cochlear implant users", *Hearing Research*, vol. 299, pp. 29-36, 2013.

[19] I. Peretz, N. Gosselin, B. Tillmann, L. Cuddy, B. Gagnon, C. Trimmer, S. Paquette and B. Bouchard,"On-line identification of congenital amusia ", *Music Perception: An Interdisciplinary Journal*, vol. 25, no. 4, pp. 331-343, 2008.

[20] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer (Version 6.0.12) [Computer program]", *Praat.org*, 2014. [Online]. Available: http://www.praat.org. [Accessed: 04-Sep- 2014].

[21] N. Macmillan and C. Creelman, *Detection theory*. Mahwah, N.J.: Lawrence Erlbaum, 2005.

[22] IBM Corp. *IBM SPSS Statistics for Macintosh, Version 20.0*. Armonk, NY: IBM Corp. Released 2011.

[23] F. Liu, C. Jiang, B. Wang, Y. Xu and A. Patel, "A music perception disorder (congenital amusia) influences speech comprehension", *Neuropsychologia*, vol. 66, pp. 111-118, 2015.

[24] P. Mok, D. Zuo and P. Wong, "Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese", *Lang Var Change*, vol. 25, no. 03, pp. 341-370, 2013.

[25] G. Peng and W. Wang, "Tone recognition of continuous Cantonese speech based on support vector machines", *Speech Communication*, vol. 45, no. 1, pp. 49-62, 2005.

[26] R. Bauer, C. Kwan-hin and C. Pak-man, "Variation and merger of the rising tones in Hong Kong Cantonese", *Lang. Var. Change*, vol. 15, no. 02, 2003.

[27] R. Fung and C. Wong, "Acoustic analysis of the new rising tone in Hong Kong Cantonese.", *In Proceedings of the 17th International Congress of Phonetic Sciences*, pp. 716-718, 2011.

[28] Y. Huo Chan and F. Yuen-yuen, *A perceptual study of tones in Cantonese*. [Hong Kong], 1974.

[29] J. Kej, V. Smyth, L. So, C. Lau and K. Capell, "Assessing the accuracy of production of Cantonese lexical tones: a comparison between perceptual judgement and an instrumental measure", *Asia Pacific Journal of Speech, Language and Hearing*, vol. 7, no. 1, pp. 25-38, 2002.

[30] S. Law, R. Fung and C. Kung, "An ERP Study of Good Production vis-à-vis Poor Perception of Tones in Cantonese: Implications for Top-Down Speech Processing", *PLoS ONE*, vol. 8, no. 1, p. e54396, 2013.

[31] J. Ou, S. Law and R. Fung, "Relationship between individual differences in speech processing and cognitive functions", *Psychonomic Bulletin & Review*, vol. 22, no. 6, pp. 1725-1732, 2015.