9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Unequal Impairment of Native and Non-native Tone Perception in Cantonese Speakers with Congenital Amusia

Gaoyuan Zhang[1], *Jing Shao*[1], *Xunan Huang*[1], *Lan* Wang[2] and *Caicai Zhang*[1, 2]

[1]Department of Chinese and Bilingual Studies, the Hong Kong Polytechnic University
[2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
gynzhang@polyu.edu.hk, jing.shao@polyu.edu.hk, 16902561r@connect.polyu.hk, lan.wang@siat.ac.cn, caicai.zhang@polyu.edu.hk

Abstract

Congenital amusia is a neurogenetic deficit that impacts pitch processing in music. Studies have shown that the deficit in amusia not only affects pitch processing in music, but also transfers to the language domain, influencing pitch processing in speech, such as lexical tone and intonation perception. Previous studies have shown that amusic individuals are impaired in lexical tone perception in both native and non-native language speakers. However, it is still unclear whether individuals with amusia are more impaired in the perception of native tones, which have long-term phonological representations, or non-native tones, which depends more on auditory/phonetic pitch processing. To fill this gap, this study examined the discrimination of pairs of native Cantonese tones and non-native Thai tones by 14 Cantonese speakers with amusia and 14 controls. Results showed that Cantonese-speaking amusic individuals were more impaired in the discrimination of non-native Thai tones than native Cantonese tones, suggesting a profound impairment in auditory/phonetic pitch processing in amusia. This finding also suggested that early exposure to a tonal language might not compensate for the impairment of lexical tone processing in a non-native language.

Index Terms: congenital amusia, lexical tone, Cantonese, Thai, tone discrimination.

1. Introduction

Congenital amusia (amusia, hereafter) is a neurodevelopmental disorder that affects musical processing throughout lifetime without brain injury [1, 2]. About 1.5-4% of the general population is affected by amusia [1, 3]. Evidence has shown that the deficit in amusia can also extend to the language domain, influencing the perception and production of lexical tones and linguistic prosody, which means that amusia is a domain-general deficit [4-8].

Concerning lexical tone perception in individuals with amusia (amusics, hereafter), both non-tonal and tonal language speakers have been investigated. For non-tonal language speakers with amusia, it has been found that French-speaking amusics performed worse than controls in the discrimination of Mandarin tones [9]. Another study has reported that French speakers with amusia have difficulties in discriminating lexical tone distinctions in Mandarin and Thai [10]. It has been suggested that amusic participants in non-tonal languages have deficits in non-native lexical tone processing.

In tonal languages, pitch is used to distinguish lexical meanings. Similar to amusic individuals in non-tonal languages, amusic individuals with tonal language background also exhibit the pitch disorder despite early exposure to speech-relevant pitch contrasts [11-16].

As for Mandarin-speaking amusic individuals, Nan et al. [11] found that in the tone discrimination task, the amusic group was not significantly impaired when tonal contrasts were carried by the same syllables, but was impaired in detecting tonal contrasts carried by different syllables. It was suggested that compared with the control group, the amusic group might have more difficulties when they had to filter out irrelevant variations such as syllables, implying possible impairment in phonological processing in amusia. In another study, Jiang et al. [12] examined the categorical perception of lexical tones in the amusic group using the tonal continua from the high-level tone to the mid-rising tone and from the high-level tone to the high-falling tone. Results indicated amusic participants were impaired in the categorical perception of native Mandarin tones. Besides, Wang and Peng [13] found that when Mandarin-speaking amusic individuals discriminated Cantonese level tones carried by familiar and unfamiliar syllables, the accuracies of both were similarly low, whereas the control group did better when tones were carried by familiar syllables than by unfamiliar Cantonese syllables. This indicated that phonological knowledge of native syllables could facilitate the control group in discriminating different tones, whereas amusics failed to show such benefit.

Apart from Mandarin-speaking amusic individuals, several studies have investigated the perception of lexical tones by Cantonese-speaking amusic individuals. For example, Shao et al. [14] found that compared with the control group, the amusic group was less accurate in the identification and discrimination of tones in all signal-to-noise ratio (SNR) conditions. The amusic participants also responded more slowly in the discrimination task. Liu et al. [15] found that for both speech and non-speech conditions, the controls achieved better performance than amusics on the discrimination of four Cantonese tone pairs (T1 high-level tone vs. T2 high-rising tone, T2 vs. T5 long-rising tone, T4 low-falling tone vs. T6 low-level tone and T5 vs. T6). Zhang et al. [16] found that compared with controls, Cantonese-speaking amusics performed less categorically in the perception of lexical tones (using the tonal continuum from T1 to T2), which suggested a deficit of the amusic group in the higher-level phonological processing of lexical tones.

Taken together, the above studies suggested that amusia is a domain-general deficit, which is not restricted to music perception. The deficit in amusia influences the perception of native as well as non-native tones. However, it is still unclear whether amusics are more impaired in the perception of native

 10.21437/SpeechProsody.2018-114

tones or non-native tones. This comparison may provide some insight into whether compared with the processing of non-native tones the established long-term phonological representation affects pitch processing in native tones. Furthermore, it is well documented that individuals with tonal language backgrounds showed more advantages than non-tonal language speakers in processing non-native lexical tones [17-19]. However whether tonal language experience may also facilitate the performance of non-native lexical tone perception in amusics is not clear. No previous studies have directly compared the perception of native and non-native tones. To this end, we tested a group of Cantonese-speaking amusics and age- and gender-matched controls on the discrimination of lexical tones in Cantonese and Thai. For lexical tones in Cantonese, established long-term phonological representations may be activated, facilitating lexical tone processing in native tones. For lexical tones in Thai, listeners have to rely on primarily auditory/phonetic pitch processing of the tones. Since amusics are reported to show deficits in both pitch processing and phonological processing in lexical tones [4-16], we expect the amusics to show worse performance than controls in both native and non-native tones.

2. Method

2.1. Participants

14 amusics and 14 controls were recruited in Shenzhen for the experiment. All of them were native Cantonese speakers who had mixed Cantonese dialect background, such as Guangzhou and Zhanjiang dialects. They were also fluent Mandarin speakers. Amusic and control participants were right-handed and matched in age and gender. None of them were reported to have any history of musical training, hearing impairments or brain injuries. None of them had knowledge of Thai. All subjects were chosen using the Online Identification Test of Congenital Amusia [20] that is composed of three parts: out of key, offbeat and mistuned tests. The cutoff score used for the selection of amusic participants in this study was 70, which was the average score of three sub-tests. A brief summary of characteristics of participants is given in Table 1.

2.2. Stimuli

To assess lexical tone discrimination in Cantonese, 12 words based on two syllables (/ji/ and /fu/) that are minimally contrastive in six lexical tones were selected as stimuli (T1: high-level tone, T2: high-rising tone, T3: mid-level tone, T4: low-falling tone, T5: low-rising tone, T6: low-level tone). A female Cantonese speaker was recorded reading aloud the selected syllables. Figure 1 shows the Cantonese speaker's F0 profile for the six Cantonese tones.

To assess lexical tone discrimination in Thai, 15 words based on three syllables (/pa/, /k^{h}a/ and /si/) carrying five lexical tones were selected (T1: mid tone, T2: low tone, T3: falling tone, T4: high tone, T5: rising tone). A female Thai speaker recorded the stimuli. Figure 2 displays the F0 profile for the five Thai tones. All the stimuli were normalized in duration to 500ms and in average intensity to 70 dB by Praat scripts.

2.3. Procedure

The study included a discrimination task, implemented in E-prime 2.0. The whole experiment was divided into two blocks: one for Cantonese tones and the other for Thai tones.

Table 1: *Demographic characteristics of Cantonese subjects.*

	Amusic Individuals	Controls
Male/Female (Total)	6/8 (14)	6/8 (14)
Mean Age (range)	20.36 (16-26)	21.58 (17-27)
Test of Congenital Amusia		
Global Score (SD)	61.78 (7.24)	86.33 (3.82)
Out of key (SD)	64.57 (11.76)	89.5 (4.95)
Offbeat (SD)	61.5 (9.20)	83.08 (8.34)
Mistuned (SD)	60.64 (8.08)	86.25 (7.14)

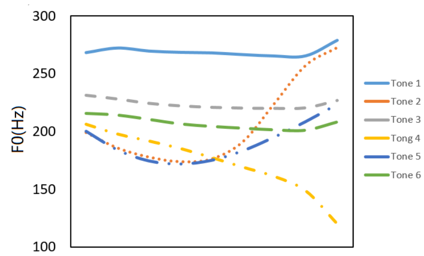

Figure 1: *The F0 profile for the six Cantonese tones.*

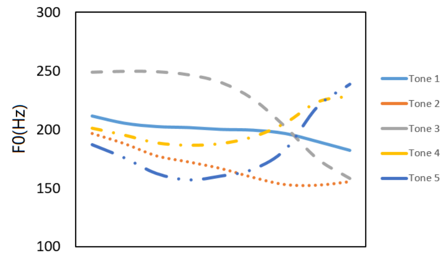

Figure 2: The *F0 profile for the five Thai tones.*

There were practice sessions before each block to familiarize participants with the experiment. Half of the subjects did the Cantonese tone discrimination block first and the other half did the Thai tone discrimination block first. In the Cantonese tone discrimination block, stimuli (/ji/ and /fu/) were presented in two sub-blocks separately. For each syllable, there were six same tone pairs and 15 different tone pairs. Within a sub-block, the same tone pairs were repeated five times and different tone pairs were repeated twice to make sure that different tone pairs and same tone pairs occurred with equal probability. All pairs were intermixed and presented randomly. In each trial, the participants were instructed to respond by pressing the "left arrow" button if the two tones were the same and the "right arrow" button if the two tones were different. A fixation first occurred on the computer screen for 500ms and two stimuli with duration of 500ms each separated by an inter-stimulus-interval of 500ms were auditorily presented. The maximal response time for subjects was 5000ms. Discrimination accuracy and response time were collected.

In the Thai tone discrimination block, tone pairs carried by the same syllable were presented in one sub-block. The same tone pairs were repeated four times while different tone pairs occurred two times. The other procedure was the same as that in the Cantonese tone discrimination block.

2.4. Data analysis

For both Cantonese and Thai tone discrimination blocks, accuracy and response time were analyzed. The accuracy of the discrimination task was calculated by the sensitivity index d' [21], which was computed as the z-score value of hit rate (the proportion of "different" responses among different tone pairs) minus false alarm rate (the proportion of "different" response among the same tone pairs). As for response time analysis, incorrect trials were disregarded). The d' scores were averaged across the tone pairs for each subject in each block. *Group* × *language* repeated measures ANOVAs were then conducted by Statistical Package for the Social Sciences (SPSS). *Group* × *tone pair* repeated measures ANOVAs were also conducted on the d' scores and response time of the Cantonese and Thai data respectively by SPSS

3. Results

Group (amusics and controls) × *language* (Cantonese and Thai) repeated measures ANOVAs were conducted on the d' scores, collapsing the tone pairs in each language. The results revealed a significant main effect of *language* (F (1, 26) = 11.02, p = 0.003) and significant *group* × *language* interaction (F (1, 26) = 7.23, p = 0.012). Independent sample t-tests were conducted to explore the significant interaction. Within each language, amusics performed significantly worse than controls in the discrimination of Thai tones (t (26) = -2.55, p = 0.017), but not in the discrimination of Cantonese tones (t (26) = -0.68, p= 0.501). Within each listener group, the amusic group performed significantly better on Cantonese tones than for Thai tones (t (26) = -2.21, p = 0.036), whereas the control group did not show such difference (t (26) = -0.38, p = 0.71).

To further analysis the data, *Group* × *tone pair* repeated measures ANOVAs were also conducted on the d' scores and response time of the Cantonese and Thai data respectively. Figure 3 shows the d' scores in the discrimination of Cantonese tones. For Cantonese tone discrimination, there was a significant main effect of *tone pair* (F (5,189, 134.918) = 12.488, p < 0.001), but there were no significant main effects of *group* or two-way interaction between *tone pair* and *group*. Post hoc analysis using the Bonferroni correction revealed that the score of the pair T2/T5 (M = 2.918, SD = 0.304) was much higher than other tone pairs (ps < 0.05), except for pairs T1/T3 (M = 4.155, SD = 0.218) and T3/T6 (M = 2.99, SD = 0.276). The score of the pair T3/T6 was also significantly higher than other tone pairs (ps < 0.05), expect for the pair T2/T5. These results suggested that the tone pair T2/T5 was most challenging to Cantonese speakers, followed by the tone pair T3/T6. In both pairs, the two tones are acoustically quite similar to each other (T2/T5: high-rising/low-rising; T3/T6: mid-level/low-level).

Figure 4 shows the mean response time in the Cantonese discrimination block. There was a significant main effect of *tone pair* (F (14, 336) = 3.49, p < 0.001), but there were no significant main effects of *group* or two-way interaction between *tone pair* and *group*. Post hoc tests showed that only the mean response time of the pair T2/T6 (M = 477.26ms，SD = 42.426) was significantly shorter than pairs T2/T5 (M = 805.195ms, SD = 69.82, p = 0.001) and T5/T6 (M = 824.295ms, SD = 93.087, p = 0.008). The results suggested it required more efforts to discriminate the tone pair T2/T5.

Figure 5 shows the d' scores in the discrimination of Thai tones. There were significant main effects of *group* (F (1, 26) = 6.486, p = 0.017) and *tone pair* (F (5, 38, 139.887) = 7.447, p < 0.001). There was no significant two-way interaction. For the main effect of *group*, the d' score of amusics (M = 3.378, SD = 1.105) was significantly lower than controls (M = 3.378, SD = 0.849). For the tone pairs, post hoc tests with the Bonferroni correction revealed that the d' score of the pair T4/T5 (M = 2.968, SD = 0.197) was significantly lower than pairs T1/T4, T1/T5, T2/T3, T2/T4, T3/T4 and T3/T5 (ps < 0.05). Furthermore, the d' score of the pair T1/T3 was significantly lower than the pair T1/T4 (p = 0.01). The results revealed that for Cantonese speakers, the tone pair T4/T5 in Thai was the most difficult to discriminate.

Figure 6 shows the mean response time in the discrimination of Thai tones. There was a significant main effect of *tone pair* (F (9, 279) = 3.242, p = 0.001), but there were no significant main effects of *group* or two-way interaction between *tone pair* and *group*. Post hoc analysis indicated that the mean response time of the pair T4/T5 (M = 735.989ms, SD = 58.992) was significantly longer than pairs T1/T5 (M = 521.471ms, SD = 43.604), (p = 0.007) and T3/T5 (M = 553.029ms, SD = 49.223), (p = 0.024), which indicated that Cantonese speakers needed longer time to discriminate the tone pair T4/T5 (high/rising tone) in Thai.

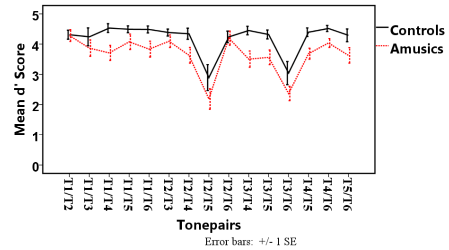

Figure 3: *The sensitivity index d' of Cantonese discrimination.*

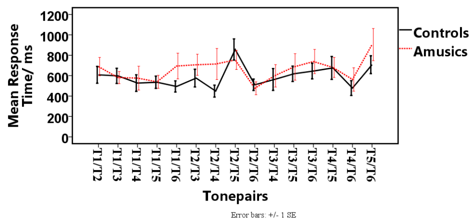

Figure 4: *The mean response time of Cantonese discrimination*.

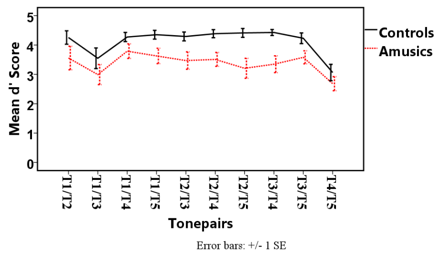

Figure 5: *The sensitivity index d' of Thai discrimination.*

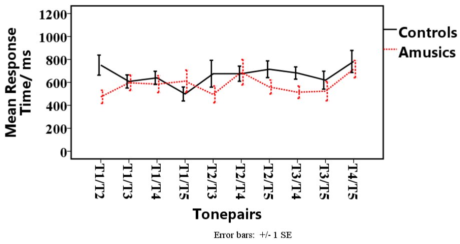

Figure 6: *The mean response time of Thai discrimination.*

4. Discussion

The present study aimed to explore the differences in the discrimination of native and non-native lexical tones in a group of tonal language speakers with amusia. 14 Cantonese-speaking amusics and 14 controls were tested on the discrimination of lexical tones in Cantonese and Thai. We found that the amusic group performed worse than controls, but not significantly, on the discrimination of Cantonese tones. However, amusic participants significantly underperformed in the discrimination of Thai tones compared to the controls. Furthermore, amusics showed better performance on Cantonese tones than for Thai tones, while the controls did not show such tendency. These findings indicated that the amusics were more impaired in the discrimination of non-native tones than native tones compared with the control group. Together with previous studies, these results suggested that the deficit of amusia is domain-general, rather than domain-specific. Moreover, early exposure to tonal language does not appear to facilitate amusics' discrimination of non-native lexical tones.

The finding that Cantonese amusics performed worse but not significantly than controls in Cantonese tone discrimination deviated from the results of Shao et al [14], where Cantonese-speaking amusics showed significantly inferior performance in tone discrimination. It might be because subjects in the previous study were Hong Kong Cantonese speakers whose language background was more homogeneous, whereas Cantonese speakers in the current study had mixed Cantonese dialect backgrounds. Nonetheless, our results were consistent with a previous study [11], where Mandarin-speaking amusics were not significantly impaired in Mandarin tone discrimination when tones were carried by the same syllables.

The main difference between Cantonese and Thai is that Cantonese is the participants' native language but Thai is a non-native language. Cantonese-speaking amusics are more familiar with Cantonese stimuli due to long-term experience with the language. But this kind of knowledge might not facilitate amusics' performance of discrimination of Cantonese tones to be as good as controls' performance because amusics still performed worse but not significantly than controls. We also found that amusics achieved significantly worse performance than controls on the discrimination of Thai tones. Subjects in the present study did not have any Thai experience, meaning that their discrimination of Thai tones primarily depends on auditory/phonetic processing of pitch distinctions. This result was consistent with previous findings [1] that amusics have severe deficits in the auditory processing of pitch differences. Importantly, we found that the amusics were more impaired in the discrimination of non-native tones than native tones compared with controls, which meant that although amusics have a deficit in phonological processing [11, 13, 16], they might be more impaired in pitch processing. It suggested that impaired phonological representations might still influence lexical tone discrimination in native language.

The finding that tonal language speakers with amusia were impaired in processing non-native tones was compatible with Wang and Peng's finding that Mandarin-speaking amusics demonstrated difficulties in discriminating non-native Cantonese tones [14]. As for non-tonal language speakers, previous studies have also found that the performance of amusics was significantly worse than the control group in the perception of non-native lexical tones [9, 10]. These consistent results suggested that amusics were inferior to controls in the discrimination of non-native lexical tones, no matter whether they have early exposure to a tonal language or not. This further suggests that early exposure to a tonal language offers no compensation for congenital amusics in the perception of non-native lexical tones.

The finding that tonal language experience has little facilitating effects on the perception of non-native tones was different from that of second-language learners. It is well documented that compared with non-tonal language speakers, individuals with tonal language backgrounds showed advantages in processing non-native tones [17-19]. For example, Scharfer et al. [19] found that the discrepancy in lexically contrastive pitch of first languages led to a hierarchy of perceptual accuracy of non-native tones. The difference between second-language learners and amusics probably results from the impairment of amusics in pitch processing, not the lack of robust phonological categories of tones in non-native languages.

In both Cantonese and Thai, certain pairs of tones were more difficult to discriminate than others. It is interesting that there is some similarity between the most challenging tone pairs in Cantonese and Thai. The result of Cantonese discrimination showed that the tone pair T2/T5 was the most difficult to discriminate, which echoed previous findings [15]. As for Thai tones, both amusic and control groups exhibited difficulties in discriminating the tone pair T4/T5. It is notable that these two tone-pairs, T2/T5 in Cantonese, and T4/T5 in Thai, both bear a rising contour (seeing Figure 1 and 2). It is possible that similar acoustic features of these tone pairs made them more difficult for the subjects to discriminate.

5. Conclusions

Results of this study showed that Cantonese speakers with amusia were more impaired in the discrimination of non-native Thai tones than native Cantonese tones. This suggested that the amusics had better performance of tone discrimination in their native language than that in a non-native language. The deficit in amusia could be better understood as a domain-general phenomenon that is not restricted to music perception. Furthermore, early exposure to a tonal language might not facilitate the discrimination of non-native tones in amusics.

6. Acknowledgements

This work was supported by grants from the Research Grants Council of Hong Kong (ECS: 25603916), the National Natural Science Foundation of China (NSFC: 11504400, 61135003， 91420301), the PolyU Start-up Fund for New Recruits (Project account no.: 1-ZE4Y).

7. References

[1] 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

[2] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proceedings of the IEEE, vol. 77, no. 2, pp. 257–286, 1989.

[3] 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.

[4] N. Gosselin, P. Jolicoeur and I. Peretz, "Impaired memory for pitch in congenital amusia", Annals of the New York Academy of Sciences, vol. 1169, pp. 270–272, 2009.

[5] B. Tillmann, K. Schulze and J. M. Foxton, "Congenital amusia: a short-term memory deficit for non-verbal, but not verbal sounds," Brain & Cognition, vol. 71, pp. 259-264, 2009.

[6] V. J. Williamson and L. Stewart, "Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem," Memory, vol. 18, pp. 657-669, 2010.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] J. Shao, C. Zhang, G. Peng, Y. Yang, and S. Y. Wang, "Effect of Noise on Lexical Tone Perception in Cantonese-Speaking Amusics," in INTERSPEECH, 2016, pp. 272-276.

[15] F. Liu, A. H. D. Chan, V. Ciocca, C. Roquet, I. Peretz, and P. C. M. Wong, "Pitch perception and production in congenital amusia: Evidence from Cantonese speakers," Journal of the Acoustical Society of America, vol. 140, p. 563, 2016.

[16] C. Zhang, J. Shao, and XN. Huang. "Deficits of congenital amusia beyond pitch: Evidence from impaired categorical perception of vowels in Cantonese-speaking congenital amusics." PloS one, vol. 12, no. 8, e0183151, 2017.

[17] A. L. Francis, V. Ciocca, L. Ma, and K. Fenn, "Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers," Journal of Phonetics, vol. 36, pp. 268-294, 2008.

[18] R. Wayland and S. Guion, "Training English and Chinese listeners to perceive Thai tones: A preliminary report." Language Learning vol. 40, no. 4, pp. 681-712, 2004.

[19] V. Schaefer and I. Darcy, "Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones," Laboratory Phonology, vol. 5, pp. 489-522, 2014.

[20] 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.

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