10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

The effects of tonal experience on the categorization of Cantonese lexical tones into Japanese native pitch accent categories

Janice Wing-Sze Wong[1], *Takayuki Arai*[2]

[1]Hong Kong Baptist University, Hong Kong SAR
[2]Sophia University, Japan
janicewong@hkbu.edu.hk, arai@sophia.ac.jp

Abstract

This study examined the effects of prosodic experience in the first (L1) and second (L2) language on the perception of non-native lexical tones. Japanese naïve listeners (Group JN) and Japanese learners of Mandarin (Group J1M2) were instructed to categorize the six Cantonese lexical tones into their native pitch accent categories. Results showed that both groups could only categorize two tones into their native pitch accent categories, and their categorization patterns were different: Group JN only assimilated two rising tones, T2 and T5, into the final-accented LH* while Group J1M2 assimilated high-rising T2 into LH* and low-falling T4 into initial-accented H*L. These preliminary results are only partly compatible with the assumptions of Perceptual Assimilation Model for Suprasegmentals (PAM-S) stating that non-native tonal categories will be assimilated to native prosodic categories, but provide more evidence that learning a more complicated system with lexical F0 variations at the phonetic level (i.e. Mandarin) does influence the perceptual assimilation of non-native tones.

Index Terms: lexical tone perception, Perceptual Assimilation Model for Suprasegmentals (PAM-S), tonal experience, Japanese pitch accent, Mandarin

1. Introduction

Lexical tones are reported to be difficult for non-tonal language learners to perceive, even though pitch may still be used in their first language (L1). In tonal languages, all words are contrastive for tone, i.e., each tone on the same syllable is associated with a different meaning. Learners acquiring a new language will have to redirect their attention to the acoustic patterns which have been ignored [1] and/or learn to ignore those which have been used in the native language but not the new one [2]. The acoustic patterns of the language per se are reported to influence the perception of tones similarly to all listeners, meaning that whether or not native or non-native listeners rely on a specific cue to distinguish the tones depends on the saliency or the particular acoustic property of the cue itself. For instance, [3] reported that, regardless of the participants' linguistic experience, acoustically easy tone pairs in Cantonese with larger acoustic differences were perceived more accurately than acoustically difficult tone pairs with smaller acoustic differences. Similarly, [4] and [5], respectively, showed that the onset of the height the fundamental frequency (F0) of Thai and Mandarin tones were more salient than their corresponding F0 offset.

Besides acoustic factors, some other tone perception studies have also shown that speakers of non-tone and tone languages may still react differently to the various dimensions of F0 in a new language, even though F0 may still be the primary cue for tone perception. This difference in perception of tones may be attributed to listeners' previous linguistic experience, i.e., their previously acquired languages, which has become a source of transfer in learning the new, unfamiliar tone language. For instance, while native speakers of tone languages perceive tones categorically, speakers of non-tone languages may judge based on the psychoacoustic properties of the stimuli instead [6]. Even native speakers of different tone languages may perform in a different way because they have a different native tonal inventory size [7] or they are used to different acoustic properties of the tones or prosodic patterns as perceptual cues in their own native language [8]. For example, Mandarin listeners are more sensitive to F0 direction than F0 height; Cantonese listeners attend to both; English listeners pay more attention to F0 slope and average F0 (see [1, 9, 10] for more information). When asked to discriminate tones in an unfamiliar language, due to the familiarity of using pitch variations in the native language, these speakers with L1 tonal language experience performed better than those without relevant experience in their L1 in general. Nevertheless, the influence of the L1 tonal experience has not been found unanimously positive, as some studies such as [11] and [12] showed that previous tonal experience could indeed hinder non-native lexical tone perception.

From the studies above, it is evident that linguistic background plays a crucial role in the perception of non-native tones. This is in line with a category assimilation account, the Perceptual Assimilation Model (PAM) [13] and PAM-L2 (i.e., the PAM for second language learners) [14], or more specifically, the Perceptual Assimilation Model for Suprasegmentals (PAM-S) [12, 15-17]. PAM-S assumes that naïve listeners will categorize non-native tones into their native L1 prosodic ones (e.g. tones, pitch accent, intonation) according to the phonetic and phonological similarities of the two languages. For instance, it is found that Mandarin tones could be categorized into Japanese pitch accent categories which are phonetically similar to the Mandarin tones [18]; Cantonese listeners categorized Mandarin tones based on the similarity of the F0 contours of the two languages [16]; English and French listeners also showed a different pattern when categorizing Mandarin tones into their intonation categories due to the difference in the use of lexical stress [19].

While an extensive number of studies have explored the effects of native linguistic experience on the perception of L2 lexical tones, little has been known for the case of non-native lexical tone perception of speakers with L2 tonal experience. One recent exception [3] investigated the reliance of F0 cues in the perception of L3 Cantonese tones by English learners of Mandarin. Their results showed that both L1 and L2 experience modulates L3 tonal perception, and if L2 and L3

10.21437/SpeechProsody.2020-99

are both tonal languages, it is L2 experience that modulates L3 tonal perception. Still, the study only focused on a discrimination experiment of two tone pairs, and did not investigate native tonal categorization, which is fundamental to the understanding of how a listener perceive non-native tones: is it constrained by the phonological and/or phonetic properties of their L1 and/or L2? To bridge the gap, the present study expanded on the aforementioned non-native tone perception research by investigating how two groups of native speakers of Tokyo Japanese (*Japanese* hereafter)—naïve listeners and those who learnt Mandarin as an L2—categorized Hong Kong Cantonese (*Cantonese* hereafter) lexical tones into their native pitch-accent categories.

These three languages all make use of F0 as one of the cues to mark the lexical tones or accented syllables. Mandarin is a tonal language with four different tones that contrast meanings: T1 (high-level, 55), T2 (rising, 25), T3 (dipping, 214), and T4 (falling, 51). While F0 is the primary cue for Mandarin tones, amplitude and duration are informative for the identification of tones even when the pitch cues are unavailable, as shown in studies using stimuli in whisper or with F0 neutralized [20]. Cantonese has six different lexical tones: T1 (high-level, 55), T2 (high-rising, 25), T3 (mid-level, 33), T4 (low-falling, 21), T5 (low-rising, 23), and T6 (low-level, 22) [21, 22]. Similar to Mandarin, Cantonese uses F0 as the primary cue for distinguishing the tones. Other acoustic cues such as duration and amplitude, however, do not appear to be useful cues for discrimination of tones [23]. Even though the low-falling tone T4 may involve some degree of glottalization or creakiness, it has not been a consistent cue even for native Cantonese listeners [24]. Thus, F0 appears to be the primary or even the only cue for lexical tone discrimination in Cantonese [25, 26]. Japanese is a non-tonal language; however, it still has a prosodic pattern known as pitch accent, using combinations of high (H) and low (L) pitch (F0) and an accent (*) to contrast meanings in words ranging from monomoraic to four-mora words [27]. For instance, for bimoraic words, *hashi* (initial-accented H*L, "chopsticks"), *hashi* (final-accented LH*, "bridge"), and *hashi* (unaccented LH, "edge") are contrastive, differing only in the pitch accent pattern. In terms of phonetic properties, the pitch accent is characterized by a higher F0 and amplitude, and longer duration in the accented syllables than the unaccented ones [28, 29]. Sugito [30] reported that Japanese pitch accent is acoustically realized by various acoustic cues, with F0 contour being the predominant cue while amplitude and duration being the weaker secondary cues. With the differences between the use of F0 cues in these three languages in mind, this present study thus allows us to establish how L1 prosodic experience and/or L2 tone learning will influence the perception of unfamiliar non-native lexical tones.

Recall that one assumption of the PAM-S is that categorization would occur if the phonetic or phonological properties of Cantonese tones are compatible with the pitch accent system of Japanese listeners. Based on the results of previous studies [e.g. 15-18], it was thus predicted that native Japanese speakers would assimilate Cantonese level tones T1, T3, and T6 into Japanese unaccented LH (or uncategorized), rising tones T2 and T5 into final-accent LH* and the low-falling T4 into initial-accent H*L. Meanwhile, if the group with L2 tonal experience demonstrated a different categorization pattern from the naïve listener group, it would suggest that an addition of tonal experience (i.e. L2 Mandarin) also has an effect on the perception of non-native Cantonese lexical tones. Note that there was also a possibility that neither group assimilated the Cantonese tones into the pitch-accent categories, i.e. they would not show any robust selection of responses; rather, a random selection will be demonstrated.

This paper is only part of a larger study that examines the effects of tonal experience on the perception of Cantonese lexical tones involving a series of experiments. Still, the results of non-native tonal categorization would deepen our understanding of whether native speakers of a pitch accent system would assimilate an unfamiliar lexical tone system into their own prosodic categories and whether an addition of L2 tonal experience would modulate the categorization. Both provide more evidence as to the role of L1 and/or L2 phonetic and phonological properties in non-native tone perception.

2. Method

2.1. Participants

Two groups of participants (N=16) took part in the experiment. Only those without any formal musical training were recruited, as studies (e.g. [31]) reported the facilitatory effects of musical training in non-native tone perception. Group JN consisted of 8 native Tokyo Japanese speakers (4F4M; mean age 21.82 years) who had no prior tonal language experience. Group J1M2 had another 8 Tokyo Japanese native speakers (6F2M; mean age 21.65 years) who learnt Mandarin as a second or foreign language for an average of 2.34 years (SD = 1.65). All participants reported no hearing or speaking deficits, and neither had they learnt nor had extensive exposure to Cantonese. They also did not have any other language experiences.

2.2. Stimulus materials

Two native speakers of Cantonese (1F1M; mean age 29 years) were invited to produce the stimuli for this experiment. The stimuli were six Cantonese tones on the syllable /fuː/ in a carrier sentence (see Table 1). The reason for choosing this syllable was two-fold: 1) it has been used and tested in previous studies [18]; 2) this syllable is highly similar to *fuu* (/ɸuː/) in Japanese which allows all three pitch accents and to /fuː/ in Mandarin which carries all four lexical tones. Each speaker produced five repetitions of the target words in a carrier phrase, from which three best tokens were selected and segmented for the experiment after two native speakers' verification of their intelligibility.

Table 1: *The six Cantonese stimuli used in the experiment.*

Tonal Categories	Tonal Contour	Glossary
T1 [55]	High-level	夫 "man"
T2 [25]	High-rising	撫 "caress"
T3 [33]	Mid-level	褲 "trousers"
T4 [21]	Low-falling	扶 "support"
T5 [23]	Low-rising	婦 "woman"
T6 [22]	Low-level	附 "attach"

2.3. Procedure

The experiment was carried out using ExperimentMFC in Praat [32]. Participants were instructed to categorize randomized individual presentations of 72 Cantonese stimuli (2 speakers × 6 tones × 3 tokens per tone × 2 repetitions) into

their native Japanese pitch accent categories, H*L, LH*, and LH, by pressing the button with the respective glossaries on the computer screen (see Table 2). With reference to previous studies such as [18], an additional button "Unknown" was also presented to the participants. Participants could choose the "Unknown" button if they deemed it impossible to categorize the stimulus into any of their native pitch accent categories. They were given 12 trials to familiarize themselves with the procedures before the experiment.

Table 2: *The three categories, the associated pitch accents and glossaries used in the experiment.*

Japanese Pitch Accent	Glossary on the screen (in kanji, romanization, and meaning)
H*L (initial-accented)	**夫**（婦）/ ***fuu*** *(fu)* "husband (and wife)"
LH* (final-accented)	（今）**風**? / *(ima)* ***fuu****?* "(modern) style?"
LH (unaccented)	**封**（筒）/ ***fuu*** *(too)* "envelope"

3. Analysis and Results

The categorization results of Groups JN and J1M2 will be presented separately. To claim that Japanese listeners could perceptually assimilate Cantonese lexical tones into their native pitch accent categories, the selected category has to be a) chosen significantly more than the chance level (25%) and b) the percentage of selection is also significantly higher than the other choices [19].

Figure 1 shows the perceptual assimilation patterns (in percentage) of Groups JN (upper) and J1M2 (lower). Individual t-tests for each tone were conducted for each group to compare the percentage means of pitch accent category against the chance level (25%). As the percentage of selection of the choice "Unknown" was occasionally over 10%, the data was also included in the analysis. The results showed that only the following categories were significantly higher than the chance level:, LH* for Tone 2 [$t(7) = 6.62$, $p < .001$] and LH* for Tone 5 [$t(7) = 3.63$, $p = .008$] for Group JN; and LH* for Tone 2 [$t(7) = 4.00$, $p = .005$] and H*L for Tone 4 [$t(7) = 3.83$, $p = .006$] for Group J1M2.

Chi-square tests were also carried out for the two groups separately to find out the association between the pitch accent categories (4 levels) and Cantonese lexical tone (6 levels). The association between the two factors was significant for both groups: JN: $\chi^2(15) = 259.37$, $p < .001$; J1M2: $\chi^2(15) = 250.22$, $p < .001$. Two separate two-way repeated-measures ANOVA between the pitch accent categories and the lexical tones showed the following results: for Group JN, only the main effect of pitch accent categories [$F(3,21) = 11.23$, $p < .001$] and tone × pitch accent categories interaction [$F(15,105) = 7.53$, $p < .001$] were significant; for Group J1M2, again only the main effect of pitch accent categories [$F(3,21) = 24.55$, $p < .001$] and tone × pitch accent categories interaction [$F(15,105) = 3.85$, $p < .001$] were robust.

For both groups, individual one-way ANOVAs were then conducted for each tone to gauge the differences between the pitch accent category. The analyses showed that for both groups, the effects of all pitch accent categories were significant for each tone. Table 3 displays the results.

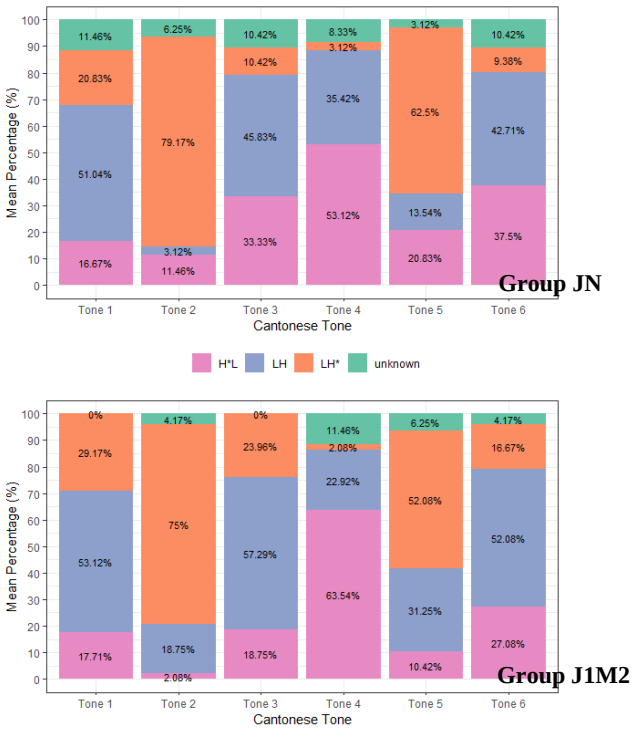

Figure 1: *Tonal categorizations (in %) for each Cantonese tone of i. (upper panel) Japanese naïve listeners [Group JN] and ii. (lower panel) Japanese-speaking learners of Mandarin [Group J1M2]. A total of 576 responses (i.e., 96 responses for each tone) was collected for each group. Asterisks indicate that the mean percentage of the choice is significantly higher than a) the chance level and b) all other categories.*

Table 3: *Results for individual one-way ANOVAs showing the effects of pitch accent categories on each tone.*

	JN	J1M2
T1	$F(3,28) = 4.50$, $p = .011$	$F(3,28) = 4.22$, $p = .014$
T2	$F(3,28) = 46.51$, $p < .001$	$F(3,28) = 8.92$, $p < .001$
T3	$F(3,28) = 8.53$, $p < .001$	$F(3,28) = 4.50$, $p = .011$
T4	$F(3,28) = 6.66$, $p = .002$	$F(3,28) = 12.26$, $p < .001$
T5	$F(3,28) = 14.41$, $p < .001$	$F(3,28) = 5.44$, $p = .004$
T6	$F(3,28) = 5.38$, $p = .001$	$F(3,28) = 6.04$, $p = .003$

Further post-hoc Bonferroni tests revealed the following results, listed in Table 4. For Group JN, only the mean percentage of assimilation of LH* category was significantly higher than the other three categories for both Tone 2 and Tone 5. For Group J1M2, only LH* and H*L were higher than each of their three counterparts for Tone 2 and Tone 4 respectively.

Table 4: *Results of post-hoc Bonferroni tests. Shaded cells indicate that the lexical tone is successfully assimilated into Japanese native pitch accent categories.*

	JN	J1M2
T1	LH>H*L (34.3%, $p = .043$) LH>UN (39.6%, $p = .014$)	LH>UN (53.1%, $p = .010$)

T2	LH*>H*L (67.7%, p <.001) LH*>LH (76%, p < .001) LH*>UN (72.9%, p< .001)	LH*>H*L (60.4%, p < .001) LH*>LH (43.8%, p = .016) LH*>UN (58.3%, p = .001)
T3	LH>LH* (35.4%, p = .002) LH>UN (35.4%, p = .002)	LH>UN (57.3%, p = .009)
T4	H*L>LH* (50.0%, p = .003) H*L>UN (44.8%, p = .010)	H*L>LH (40.6%, p = .005) H*L>LH* (61.5%, p < .001) H*L>UN (52.1%, p < .001)
T5	LH*>H*L (41.7%, p = .001) LH*>LH (49.0%, p < .001) LH*>UN (59.3%, p < .001)	LH*>H*L (41.7%, p = .018) LH*>UN (45.8%, p = .008)
T6	LH>LH* (33.3%, p = .026) LH>UN (32.3%, p = .033)	LH>LH* (35.4%, p = .032) LH>UN (47.9%, p = .002)

*UN = unknown option

4. Discussion

This study started as a preliminary investigation on the perception of Cantonese lexical tones by Japanese naïve listeners who had no other tonal experience and another group of Japanese speakers who learnt Mandarin as a second language. By observing the categorization patterns, we can obtain extra evidence to evaluate the predictions and assumptions of tonal assimilation in PAM-S, and explore further if an addition of tonal experience will influence the categorization.

The present findings, however, only partly corroborated previous studies, for not all Cantonese tones were found categorizable [13, 16, 18]. Specifically, all three level tones (T1, T3 and T6) were found uncategorized by both groups. Group JN only found the two rising tones (T2 and T5) assimilable, and they categorized the two tones into the same final-accented LH* category. Group J1M2 categorized only high-rising T2 into LH*, although they could also assimilate the low-falling T4 into initial-accented H*L. These categorization and non-categorization patterns suggested that Japanese speakers assimilated only some phonetic properties of Cantonese tones into their native pitch accent categories which share similar phonetic features, as predicted by PAM-S.

It is rather surprising to see that both groups of participants did not find the three level tones assimilable. One justification may be that the Japanese unaccented pattern may not be as level and stable as the Cantonese level tones, for the unaccented pattern itself shows a slight rise of pitch from L to H. This may deter the participants from unanimously making a confident choice. Moreover, for bimoraic Japanese words, unaccented pattern is not common at all (15%) while the initial-accent is widely found (65%) [33]. The uncategorized situation may thus be attributed to this unfamiliarity of unaccented pattern in bimoraic words. Still, there was a tendency for both groups to select the unaccented LH category (ranging from 42.71% to 57.29%, although all were not found significantly higher than the chance level or all other categories) for all level tones, signifying that the unaccented LH may already be the closest answer possible for the assimilation of Cantonese level tones. Data is being collected in the larger project to further evaluate the above speculations.

While the rising and falling tones of Cantonese display more salient difference between F0 onset and offset than the level tones, Group JN only found T2 and T5 which differ in the magnitude of F0 change in the rising contour, assimilable. Based on PAM-S, both T2 and T5 which were categorized as LH* will be interpreted as a *Single Category* (SC) assimilation pair. This was predicted to be a pair that will result in poor discrimination. This pair is found easily confusable [34], is acquired last [35], and is a potentially merging pair [36] even for native speakers of Cantonese. It would be interesting to further examine the results in the discrimination test to learn more about the perceptual patterns difficulties of such a complicated tone system by Japanese naïve listeners.

It is more intriguing to see that Group J1M2 performed slightly differently from Group JN, suggesting that tonal experience in Mandarin has strengthened the sensitivity of tone contour which in turn influenced the native categorization of the Japanese speakers. Specifically, Cantonese T2, but not T5, was categorized into LH*; T5 was found being categorized into LH* (52.08%) and LH (31.25%) with non-significant differences. The results appeared to suggest that these two tones were deemed different perceptually, and the participants might have become more sensitive to the F0 contour change. T2 has higher degree of change in the rising contour, and thus could be easily categorized as a final-accented LH*; T5 only has a minimal rise resulting in the selection of the two competing LH* and LH choices. Moreover, Group J1M2 also became more sensitive towards the low-falling tone T4 which was assimilated to H*L. It is possible that the exposure to the four L2 Mandarin tones had positively modulate the use of F0 cues in non-native Cantonese tone perception.

One limitation of the current study is that data was only from a total of 16 speakers, resulting in only 96 responses per tone in each group. Acknowledging that the sampling size might have an effect on the statistical power, we are collecting additional data to further consolidate the claims made above. Still, based on the present results, these two groups of Japanese speakers were found to be sensitive to the F0 contour of Cantonese tones, although to different degrees.

5. Concluding Remarks

This preliminary study presented the results of the perceptual assimilation patterns of Cantonese lexical tones into Japanese native pitch accent categories by Japanese naïve listeners and Japanese learners of Mandarin, hoping to further our understanding not only of the assumptions in PAM-S, but also whether prior tonal language experience influences the perception of non-native prosody in a third language. The current results did not fully support the assumptions in PAM-S that non-native prosodic categories can be assimilated to their native categories, for only two tones were successfully categorized among the Japanese naïve listener group. L2 tonal experience seemed to show some additional advantage for the assimilation as Group J1M2 was more sensitive to pitch contour than the naïve listener group. Collection of extra data from more Japanese participants and two other groups (Mandarin and Cantonese native) for benchmark comparison is underway to clarify and affirm the conclusion.

6. Acknowledgement

This work is supported by STEC Sophia Lecturing – Research Grant 2019 from Sophia University and Faculty Research Grant (FRG1/17-18/025) from Hong Kong Baptist University. The authors thank Eri Iwagami and Shujie Yang especially for their kind assistance in data collection.

7. References

[1] S. G., Guion and E. Pederson, "Investigating the role of attention in phonetic learning," in *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam, The Netherlands: John Benjamins, 2007. pp. 57–77.

[2] R. A. Yamada and Y. Tohkura, "Perception of American English /r/ and /l/ by native speakers of Japanese," in *Speech perception, production, and linguistic structure*. Tokyo: Ohmsha, 1992. pp. 155–174.

[3] Z. Qin and A. Jongman, "Does second language experience modulate perception of tones in a third language?" *Language and Speech*, vol. 59, no. 3, pp. 318–338, 2015.

[4] D. Burnham and E. Francis, E. "The role of linguistic experience in the perception of Thai tones," in *South East Asian Linguistic Studies in Honour of Vichin Panupong, Science of Language*, vol. 8, pp. 29–47, 1997.

[5] C.-Y. Lee, L., Tao and Z. S. Bond, "Identification of acoustically modified Mandarin tones by native listeners," *Journal of Phonetics*, vol. 36, no. 5, pp. 537–563, 2008.

[6] W. S.-Y. Wang, "Language change," *Annals of the New York Academy of Sciences*, vol. 28, pp. 61–72, 1976.

[7] W.-H. Chiao, B. Kabak, and B. Braun, "When more is less: Non-native perception of level tone contrasts," *Proceedings of Psycholinguistic Representation of Tone Conference*, Hong Kong, China, pp. 42–45, 2011.

[8] A. Reid, D. Burnham, B. Kasisopa, R. Reilly, V. Attina, N. X. Rattanasone, and C. T. Best, Perceptual assimilation of lexical tone: the roles of language experience and visual information," *Attention, perception & psychophysics*, vol. 77, no. 2, pp. 571–591, 2015.

[9] J. Gandour, "Perceptual dimensions of tone: Evidence from Cantonese*," Journal of Chinese Linguistics*, vol. 9, pp. 20–36, 1981.

[10] A. L. Francis, V. Ciocca, L. Ma, and Fenn, K. Fenn, "Perceptual learning of Cantonese lexical tones by tone and non-tone language listeners," *Journal of Phonetics*, vol. 36, no. 2, pp. 268–294, 2008.

[11] C. K. So. Effect of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English. PhD dissertation, Simon Fraser University, 2006.

[12] C. K. So and Best, C. T. Best, "Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences," *Language and Speech*, vol. 53, pp. 273–293, 2010a.

[13] C. T. Best, "A direct realist view of cross-language speech perception," in *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 1995, pp. 171–232.

[14] C. T. Best and M. D. Tyler, "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins, 2007. pp. 13–34.

[15] C. K. So and C. T. Best, "Do English speakers assimilate Mandarin tones to English prosodic categories?" in *Proceedings of the 9th annual conference of the International Speech Communication Association (Interspeech 2008), incorporating the 12th Australasian International Conference on Speech Science and Technology (SST 2008)*, pp. 1120–1125, 2008.

[16] C. K. So and C. T. Best, "Discrimination and categorization of Mandarin tones by Cantonese speakers: The role of native phonological and phonetic properties," in *the Thirteenth Australasian International Conference on Speech Science and Technology*, Melbourne, Australia, 2010b.

[17] C. K. So and C. T. Best, "Categorizing Mandarin tones into listeners' native prosodic categories: The role of phonetic properties," in *Poznań Studies in Contemporary Linguistics*, vol. 47, pp. 133–145, 2011.

[18] C. K. So, "Categorizing Mandarin Tones into Japanese Pitch-accent Categories: The Role of Phonetic Properties," in *Proceedings of Interspeech 2010 Satellite Workshop on Second Language Studies* (Tokyo).

[19] C. K. So and C. T. Best, "Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories," *Studies in Second Language Acquisition*, vol. 36, pp. 195–221, 2014.

[20] S. Liu and A. G. Samuel, "Perception of Mandarin lexical tones when F0 information is neutralized," in *Language and Speech*, vol. 47, no. 2, pp. 109–138, 2004.

[21] R. S. Bauer and P. K. Benedict, *Modern Cantonese phonology*. New York, NY: Mouton de Gruyter, 1997.

[22] Y. R. Chao, *Cantonese primer*. Cambridge, UK: Cambridge University Press, 1947.

[23] E. Khouw and Ciocca, V. "Perceptual correlates of Cantonese tones," *Journal of Phonetics*, vol. 35, no. 1, pp. 104–117, 2007.

[24] T. J. Vance, "An experimental investigation of tone and intonation in Cantonese," *Phonetica*, vol. 33, no. 5, pp. 368–392, 1976.

[25] V. Ciocca, A. L. Francis, R. Aisha and L. Wong, "The perception of Cantonese lexical tones by prelingually deaf cochlear implantees," *Journal of the Acoustical Society of America*, vol. 111, no. 5, pp. 2250–2256, 2002.

[26] K. Y. Lee, C. A. van Hasselt, S. N. Chiu and D. M. Cheung, "Cantonese tone perception ability of cochlear implant children in comparison with normal-hearing children," *International Journal of Pediatric Otorhinolaryngology*, vol. 63, no. 2, pp. 137–147, 2002.

[27] X. Wu, J.-Y. Tu, and Y. Wang, "Native and nonnative processing of Japanese pitch accent," *Applied Psycholinguistics*, vol. 33, pp. 623–641, 2012.

[28] M. E. Beckman, *Stress and non-stress accent*. Dordrecht, Holland: Foris, 1986.

[29] M. Sugito, "Nihongo no akusento to intoneeshon – Tokyo hogen no 'hana' to 'hana' no soui [Japanese accent and intonation – the difference between 'hana (flower)' and 'hana (nose)' in the Tokyo dialect]," in *"hana" to "hana" Nihongo no kenkyu 5. Osaka: Izumishoin. [Reprinted from Kotoba to onsei, "kotoba" Series 18]*, pp. 23–37, 1983.

[30] Y. Sugiyama, "Perception of Japanese pitch accent without F0," *Phonetica*, vol. 74, pp. 107–123, 2017.

[31] D. R. Burnham, Brooker, and A. Reid, "The effects of absolute pitch ability and musical training on lexical tone perception," *Psychology of Music*, vol. 43, no. 6, pp. 881–897, 2015.

[32] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer. Version 6.0.14", http://www.praat.org/, 2018.

[33] S. Tanaka and H. Kubozono, *Introduction to Japanese Pronunciation: Theory and Practice*. Tokyo: Kurosio, 1999.

[34] T. J. Vance, "Tonal distinctions in Cantonese," *Phonetica*, vol. 34, pp. 93–107, 1977.

[35] V. Ciocca and J. Lui, "The development of lexical tone perception in Cantonese," *Journal of Multilingual Communication Disorders*, vol. 1, pp. 141–147, 2003.

[36] P. P. K. Mok, D. Zuo, and P. W. Y. Wong, "Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese," Language Variation and Change, vol. 25, pp. 341–370, 2013.