

PERCEPTUAL ASSIMILATION EFFECTS ON L2 TONAL PROCESSING*

Jung-yueh Tu
National Chengchi University

ABSTRACT

The goal of this study is to expand our understanding of cross-linguistic comparison in non-native speech perception, focusing on perceptual assimilation at the suprasegmental level. The current study investigates perception of Mandarin tones by Cantonese listeners. In Experiment 1, Mandarin tones were presented and the Cantonese listeners were requested to identify which tone they heard. In Experiment 2, the Cantonese listeners were instructed to rate how similar each Mandarin tone was to a Cantonese tone on a 1 to 5 scale subjectively. The results suggest that tonal confusion errors may result from not only the similar acoustic properties of tone pairs but also perceptual assimilation between L1 and L2 tonal contrasts. Tone 1 and Tone 4 in Mandarin were both rated as the most similar to Tone 1 in Cantonese. Tone 2 and tone 3 in Mandarin were perceived as belonging to the overlapping set of Cantonese tonemes. The findings are discussed in terms of the effects of the L1 prosodic system on L2 perception. Additionally, the discussion explores how perceptually assimilation patterns can predict listeners' perception performance in the domain of lexical tones.

Keywords: perceptual assimilation, Mandarin tones, tonal processing, L2 perception

* This study was funded by National Science and Technology Council, Taiwan (Grant No. MOST 110-2410-H-004-046). I am grateful to three anonymous reviewers for their insightful comments. I also thank Ming-da Wu at National Cheng Kung University for his help with data collection.

1. INTRODUCTION

It is generally believed that many aspects of L2 speech learning can be explored in terms of phonological and phonetic factors which involve the similarities and differences between the L1 and L2 sound systems (Munro & Bohn 2007). A number of studies on L2 perception have paid attention to segments, such as vowels (e.g., Flege, Bohn & Jang 1997; Walley & Flege 1998) and consonants (e.g., Flege 1991; Flege, Takagi, & Mann 1995; Tsukada 2005; McMurray et al., 2002, 2009). These cross-linguistic studies have been devoted to the two most influential models, Flege's (1995) Speech Learning Model and Best's (1995) Perceptual Assimilation Model, which both highlight the notion of interference by phonetic similarity between L1 and L2 inputs at a segmental level. Suprasegments, however, are also important in L2 perception but are somewhat understudied within these theoretical frameworks. With this background, what interests us most is whether and how the similarity between L1 and L2 prosodic systems influences L2 perception. This study is designed to investigate if patterns of assimilation in prosody can also predict L2 perception. To test if one or both of these models can be applicable to the suprasegmental level, the tonal systems in two tone languages, Mandarin and Cantonese, are examined in this paper. Considering that tonal judgment is likely influenced by factors beyond acoustic properties, such as phonological awareness and perceptual assimilation, it becomes important to take these factors into account when studying tone perception.

1.1 Research Questions

The goal of this study is to investigate how the Cantonese tonal system influences the perception of Mandarin tones. The acoustic properties of the tonal systems in the two languages are compared. In addition, this study aims to investigate the error patterns exhibited by Cantonese listeners when perceiving Mandarin tones. More specifically, the study intends to answer the following question: What are the effects of the Cantonese tonal system on the perception of Mandarin tones by Cantonese listeners? In line with this question, three secondary questions arise: 1) Which Mandarin tones are more difficult for Cantonese listeners to

perceive? 2) How is each tone in Mandarin perceptually assimilated to a Cantonese tone by listeners' judgments? 3) Will tonal assimilation from Mandarin tones to Cantonese tones facilitate or interfere with the perception of Mandarin tones?

1.2 Tone Inventories of Mandarin and Cantonese

The target language of this study is Mandarin Chinese. As a tone language, Mandarin is typically described as having four tone categories: a high level tone (Tone 1) [55], a mid- rising tone (Tone 2) [35], a low-dipping tone (Tone 3) [214] and a high-falling tone (Tone 4) [51] (Chao 1930; Li & Thompson 1989; Norman 1988), as shown in Table 1. It is noted that the rising tone (tone 2) shows a dipping at the initial portion and then a moderately rising pattern in the final position (Fon & Chiang 1999).

Table 1: Tonal system in Mandarin Chinese.

Tone categories	Tonal feature	Pitch value	Height
1	high-level (HH)	55	High
2	mid-rising (MH)	35	
3	mid-level (MLH)	214	Mid
4	low-falling (HL)	51	Low

The Cantonese tonal system is considerably more complex than that of Mandarin Chinese. In Cantonese, there are six contrastive tones on non-checked syllables: a high level tone (Tone 1) [55], a mid- rising tone (Tone 2) [35], a mid-level tone (Tone 3) [33], a low-falling tone (Tone 4) [21], a low-rising tone (Tone 5) [13] and a low-falling tone (Tone 6) [22] (Bauer & Benedict 1997), as shown in Table 2. The high level tone (Tone1) [55] in Cantonese has a free variant as a high-falling tone [53], that is, a high level tone [55] and a high-falling tone [53] are in Cantonese tonal system at the phonetic level, but they are allotones of a Cantonese toneme (Vance, 1984; Bauer & Benedict 1997; Chen, 2000; So 2005). In addition, it is

reported that the two rising tones, Tone 2 [35] and Tone 5 [13], are merging in Hong Kong Cantonese (Mok & Wong, 2010).

Table 2. Tonal system in Hong Kong Cantonese.

Tone categories	Tonal feature	Pitch value	Height
1	high-level	55	High
2	mid-rising	35	
3	mid-level	33	Mid
4	low falling	21	Low
5	low rising	13	
6	low level	22	

To compare the two languages' tonal system, the pitch contours for the four Mandarin tones and six Cantonese tones in isolation, produced by native speakers, are given in Fig. 1 (adapted from Francis et al., 2008, p.271-272). (Note that in Francis et al. (2008), Tone 2 was transcribed as [25], Tone 5 as [23]).

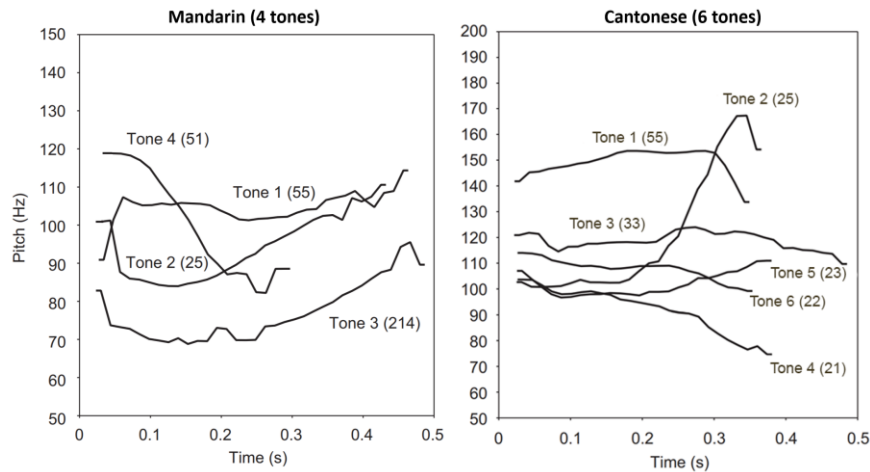


Fig.1. The pitch contours for the two languages' tones in isolation (adapted from Francis et al., 2008, p.271-272).

1.3 Background

Previous studies have examined the role of experience on adults' production and perception of L2 sounds (e.g., Flege et al. 1995; Gandour et al., 1998; Kuhl 2000; Gandour et al., 2000; So, 2005; Jongman et al., 2006; Ding, 2012). Much cross-linguistic research on consonants and vowels (see e.g., Best & Strange 1992; Flege, Takagi & Mann 1995; Cutler & Otake 1994) has suggested that native language background affects perception and production of non-native sounds at the segmental level. In line with this, Pallier et al. (1997) mentioned that people perceive speech through the filter of their native language. Regarding L2 speech perception, there are two existing most influential theoretical constructs in this field, one is Best's (1995) Perceptual Assimilation Model (PAM) and the other is Flege's (1995) Speech Learning Model (SLM) (Munro & Bohn 2007). Both the PAM and the SLM regard phonetic similarity of L2 and L1 segments as key in predicting relative perceptual difficulty for L2 learners (Strange 2007). Best and Tyler (2007) point out that the major difference between the PAM and the SLM is that the PAM has focused

primarily on naïve listeners, whereas the SLM have focused on experienced listeners of L2 phonetic acquisition over time.

PAM considers both phonological and phonetic levels in explaining the effect of the L1 system on the perception of non-native phones that are completely unfamiliar to the listeners (Best & Tyler 2007). In predicting how naïve listeners will identify and discriminate non-native phonological contrasts, PAM takes into account how a non-native phone is perceptually assimilated. A non-native phone may be perceptually assimilated as a good or even a poor exemplar of a native phonological segment (Categorized), or as unlike any single native phoneme (Uncategorized), or as a non-linguistic non-speech sound (Non-assimilated). The assimilation pattern predicts perception performance. Two Category (TC) assimilation, in which the two non-native phones are perceptually assimilated as to two different native phonemes, predicts successful discrimination. Single Category (SC) assimilation, in which the two non-native phones are heard as equally good or poor exemplars of the same native phoneme, predicts poor discrimination. Category Goodness (CG) difference, in which one of the non-native phones assimilates to a single native phoneme better than the other, predicts intermediate discrimination. Although PAM originally focused on speech segments, Halle, Chang, and Best (2004) suggested that PAM can be applicable to suprasegments as well.

As for the processing of suprasegmental features, it has been shown that the perceptual weights of the two dimensions, F0 height and contour were related to the linguistic experience of the listeners. Lee, Vakoč, and Wurm (1996) found that tone language speakers were better at discriminating L2 tones than were non-tone language speakers, in terms of speed and accuracy of their responses. Tone language speakers seemed to have acquired general tone discrimination abilities. Thus, it appears that listeners' strategy for tone perception depends to some extent on the linguistic function of pitch in their native language.

It has been found in previous prosodic perception studies that L1 suprasegmental systems have effects on the L2 perception. Some studies have found that language experience affects the tonal perception mechanism. For example, English speakers tend to be more sensitive to pitch height while Mandarin and Cantonese speakers tend to be sensitive to/aware of both pitch height and direction (e.g., Bent, 2005;

Chandrasekaran et al., 2010; Francis et al., 2008; Gandour, 1983; Gandour, 1984; Guion and Pederson, 2007). For example, Gandour (1983) used multidimensional scaling to examine the perception of tones by listeners of four tonal languages, including Mandarin, Cantonese, Taiwanese, and Thai, as well as by those of a non-tonal language, English. The results showed that English listeners attached more importance to the height and less to the contour dimension of F0 than did listeners of most tonal languages. Stagger and Downs (1993) found that speakers of tone languages are not sensitive to slight frequency changes because they make more categorical judgments of pitch. They further claim that tone language speakers tend to have more categorical perception of pitch patterns than non-tone language speakers. Halle, Chang, and Best (2004) tested French and Taiwanese listeners on discrimination of three Mandarin lexical tone continua. They found that French listeners were uniformly sensitive to f0 differences across the continua whereas Taiwanese listeners were more sensitive to differences between tokens in different categories than similar acoustic differences between pairs within same category. These studies indicate that tone and non-tone listeners may attach importance to different perceptual cues.

Several recent cross-linguistic studies on Mandarin tone perception have found that the fact certain tones (Tone 1-Tone4 pair and Tone 2-Tone 3 pair) are phonetically similar to one another in nature causes perceptual difficulties for non-native listeners (Miracle 1989; Shen 1989), and that listeners with different L1 language backgrounds showed different tonal confusion patterns (So 2005; Hao, 2012; So & Best, 2010; Tu et al., 2014, 2016; Shen et al., 2013; Qin et al., 2019). Chen (1997) found that English-speaking learners often confused Tone 2 [35] and Tone 3 [214] in perception. Wang, Spence, Jongman, and Sereno (1999) used an auditory training procedure to train American listeners to identify the four Mandarin tones. It was found that the trainees exhibited an increase in the identification of the four tones. Their tone pair confusion patterns showed that the four tones were indeed differentially acquired. Tone pair 1 and 4 was most resistant to improvement and the two tones were also found difficult for Americans to acquire (Shen 1989). Wang et al. (1999) speculated that it is probably because tone 1 and tone 4 are both comparable to the English unmarked or stressed condition, whereas the

other tone pairs each involve at least one tone that is novel or “unnatural (p.3656)” in English. It was also found that tone 2 and tone 3 were improved greatly after training, though tone 2 and tone 3 are difficult to distinguish due to their acoustic similarities. It might be that these two tones are so novel to the American listeners so that the listeners were more attentive to their distinctions. So (2005) examined perception of Mandarin tones by Cantonese and Japanese naïve listeners. Her study employed a pretest-posttest paradigm. The results revealed that errors in all the tests were mostly found in T1-T4 and T2-T3 pairs (bidirectional), except that Japanese listeners showed more error types in the pretest. Cantonese listeners produced substantially more errors in both tone pairs.

With regard to second language production, foreign accents in L2 production are caused by interference from the phonological system and phonetic realization of the speaker’s first language, including both segmental and prosodic features (See e.g. Jun & Oh 2000; Flege 1993; Flege et al. 1995). Previous research on Mandarin tone production indicates that there is no agreement on which tones are most difficult for L2 learners. Miracle (1989) found that American learners made roughly the same amount of errors across the four tones. Shen (1989) found that American learners made more errors with Tone 1 [55] and Tone 4 [51] and fewer errors with Tone 2 and Tone 3. Chen (1997) found alien level tones like [22] and [33] often replaced the target tones. These studies imply that there may be discrepancy between perception and production.

1.4 The Current Study

While a number of previous studies on L2 speech perception at the segmental level (see, e.g.; Best & Strange 1992; Flege, Takagi & Mann 1995; Cutler & Otake 1994) have been carried out under the theoretical framework of the PAM and the SLM, there are only a few studies (So 2005; Francis, Ciocca, & Fenn 2008) which have focused on perceptual assimilation at the suprasegmental level. Although So (2005) and Francis, Ciocca, Ma, and Fenn’s (2008) studies have made generalizations about tonal similarities between Mandarin and Cantonese, no studies of cross-language mapping involving overt decisions by listeners on the similarity

or goodness of fit between L2 and L1 categories have been conducted to determine the actual pattern of assimilation.

The current study proposes to compare the prosodic systems of the two studied languages in terms of acoustic properties in nature as well as perceptual assimilation by listeners' similarity ratings and further explore the tonal confusion patterns in Mandarin by native speakers of Cantonese. In particular, the goal of this study is to see how the naïve and experienced Cantonese listeners perceive Mandarin tones, which are all naturally-produced tones instead of synthesized tonal continuum. This study is designed to examine their subjective degree-of-similarity between Mandarin tones and Cantonese tones. That is, it intends to explore how similar the Cantonese speakers rate between the tones in their target language and native language influence their perceptual identification. Building upon the previous classical identification task with a focus on categorical perception, this research intends to provide an alternative perspective and additional insights to test Best' PAM.

In order to learn more about the effects of the L1 suprasegmental system on L2 learning, the prosodic systems in the two typical tone languages were compared. Specifically, according to previous findings and relevant models, this study proposes some hypotheses as follows (as summarized in Table 3):

1) Tone 1 [55] is the easiest to identify since it is categorized as Tone 1 [55] in Cantonese. Tone 4 [51] is also easier to identify compared to Tone 2 and Tone 3.

2) Tone 2 [35] and Tone 3 [214] are more difficult to perceive since they are uncategorized for Cantonese listeners. Tone 2 [35] falls between the two rising tones, Tone 2 [35] and Tone 5 [13] in Cantonese; Tone 3 [214] falls between Tone 2 [35], Tone 4 [21], and Tone 5 [13] in Cantonese.

3) As for the tonal confusion patterns, more tone errors in the tone pairs Tone 1-Tone 4 and Tone 2-Tone-3 (bidirectional) are expected (So, 2005). In terms of acoustic properties, Tone 1-Tone 4 both start at a high level and Tone 2- Tone 3 both have a rising portion of the pitch contours. Thus, the tone pairs Tone1-Tone 4 and Tone2-Tone 3 are difficult to perceive in nature (So, 2005).

4) In terms of L1 interference, Mandarin Tone 1 [55] can be categorized as Cantonese Tone 1 [55], but the free variant with a falling tone [53] of Cantonese Tone 1 may lead to difficulties in distinguishing Tone 1 [55] and Tone 4 [51] in Mandarin.

5) Also, the two rising tones in Cantonese, Tone 2 [35] and Tone 5 [13], may interfere the perception of Tone 2 [35] and Tone 3 [214] in Mandarin.

Table 3. The current hypotheses.

Hypotheses		Tonal identification	
	Mandarin tone	Similar to Cantonese tone	Predictions
1			
	Tone 1[55]	Tone 1[55/53]	easiest to identify
	Tone 4[51]	Tone 1[55/53]	easy to identify
2	Tone 2[35]	Tone 2[35] Tone 5[13]	difficult to identify
		Tone 2[35]	
	Tone 3[214]	Tone 4[21] Tone 5[13]	difficult to identify

Hypotheses	Tonal confusion patterns	
	L2 tonal features and L1 interference	Predictions
3	Mandarin Tone 1[55]-Tone 4[51] both start at a high level Mandarin Tone 2[35]-Tone 3[214] both have a rising portion	errors in T1-T4 and T2-T3 pairs
4	Mandarin Tone 1[55] and Tone 4[51] may be categorized as Cantonese Tone 1[55/53]	errors in T1-T4
5	Cantonese Tone 2[35] and Tone 5[13] may interfere the perception of Mandarin Tone 2[35] and Tone 3[214]	errors in T2-T3

2. MATERIALS AND METHODS

2.1 Experiment 1: Identification Task

A total of sixty young adults participated in Experiment 1. They were divided into two groups. The first group included 30 native Cantonese listeners (15 males and 15 females), who self-reported that they have never received formal instruction of Mandarin. The second group included 30 native Cantonese listeners (12 males and 18 females), who are learners of Mandarin. They reported that their length of learning has been more than 3 years. The Cantonese listeners all had exposure to Mandarin Chinese. All of the participants aged from 18 to 25 (mean age = 21.6). They all speak Hong Kong Cantonese as their primary dialect. None of the participants had any linguistic training or more than five years' musical training. All reported normal speech and hearing.

The stimuli were four sets of Mandarin monosyllabic words bi, liŋ, wei, and fən with four different tones (see Appendix A). A total of 16 words were used in this study. The stimuli were recorded by a female native Mandarin speaker. The presentation of stimuli was randomized. Each word was repeated two times and thus there were 32 trials (4 sets * 4 words * 2 repetitions) in total.

Before the identification task, all the listeners received a short training section (about 5 minutes) to familiarize them with the four different tones in Mandarin. An explanation of the tone patterns in terms of the pitch contour and pitch height plus visual aids was given to the trainees. They then listened to the four words recorded by the same Mandarin speaker and feedback was given after each word. Later, the listeners participated in the identification section, in which they were requested to identify the four different tones in Mandarin. The inter-trial interval (ITI) is self-paced. When a response was made, the next trial would be played in 500 ms.

2.2 Experiment 2: Similarity Rating Task

The same participants in Experiment 1 participated in Experiment 2.

The stimuli include two sets of four monosyllabic Mandarin words and six monosyllabic Cantonese words. The Mandarin words were recorded by the same speaker for Experiment 1. The Cantonese words were recorded by a female native speaker of Cantonese. All of the words in each set have the same or similar segments but differ only in tones. The syllables [fən] and [fu] were used. The two sets of Mandarin words [fən] and [fu] with four contrastive tones and the two sets of Cantonese words [fən] and [fu] with six contrastive tones were used (see Appendix B). The pitch contours of the stimuli [fu] with tones in the two languages were given in Fig. 2.

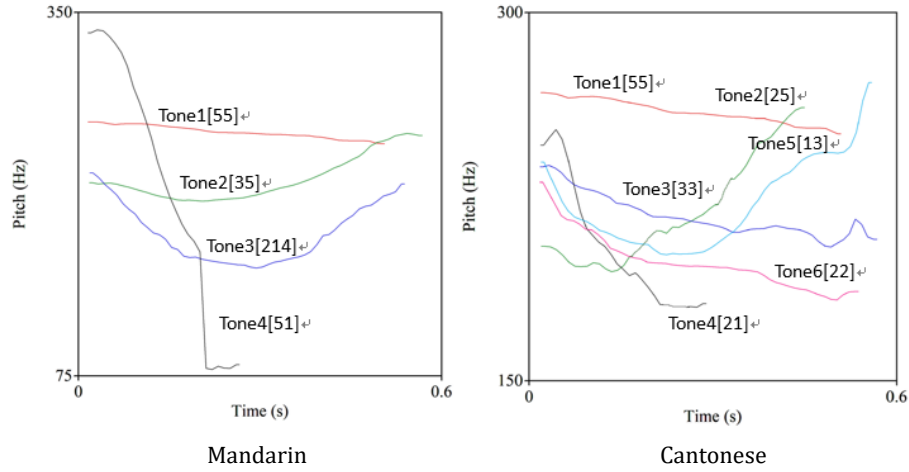


Fig. 2. The pitch contours of the stimuli [fu] with Mandarin four tones and Cantonese six tones.

It should be noted that though the vowel of the syllable *fan* [fən] in Cantonese is somewhat different from that in Mandarin [fən], the syllable was used since it can carry the six non-checked Cantonese tones and also the focus of this study is tone rather than vowel. Each Mandarin word was paired with a Cantonese word and there were two sets of 24 pairs in total. In addition, in the rating task, there may be some possible confounds, such as frequency, neighborhood density, and so on. Hence, the current study used monotones as the stimuli (without showing the corresponding Chinese characters), which may be associated with varied Chinese characters.

The participants heard 48 pairs of naturally recorded tones carried by [fən/fən] and [fu] in four sections. They were requested to rate the degree of similarity between a pair on a “1” (very different) to “5” (very similar) scale subjectively (cf. Huang, 2007). The participants were instructed to focus on the similarities of Mandarin tones and Cantonese tones according to their judgments. The inter-stimuli interval (ISI) was 300 ms, and the inter-trial interval (ITI) was self-paced. When participants responded to a

trial, the next trial would be played in 500 ms. Each participant was presented with all 48 pairs in randomized order.

3. RESULTS

3.1. Results of Experiment 1

In Experiment 1, with group and tone as independent variables, a 2 (group) \times 4 (tone) mixed-design ANOVA was applied to the tone accuracy. The main effect of the group variable was statistically significant [$F(1, 58) = 331.78$, $MSE = 119.83$, $p < .001$, $\eta^2 = .851$]. The main effect of the tone variable was also statistically significant [$F(3, 174) = 57.98$, $MSE = 101.90$, $p < .001$, $\eta^2 = .500$]. There was a statistically significant interaction between the group and tone variables [$F(3, 174) = 11.15$, $MSE = 101.90$, $p < .001$, $\eta^2 = .161$]. Further testing of the simple-main effect showed that all tone accuracy of the Learner group was significantly higher than that of the Naïve group (all p values $< .001$) (see Fig. 3).

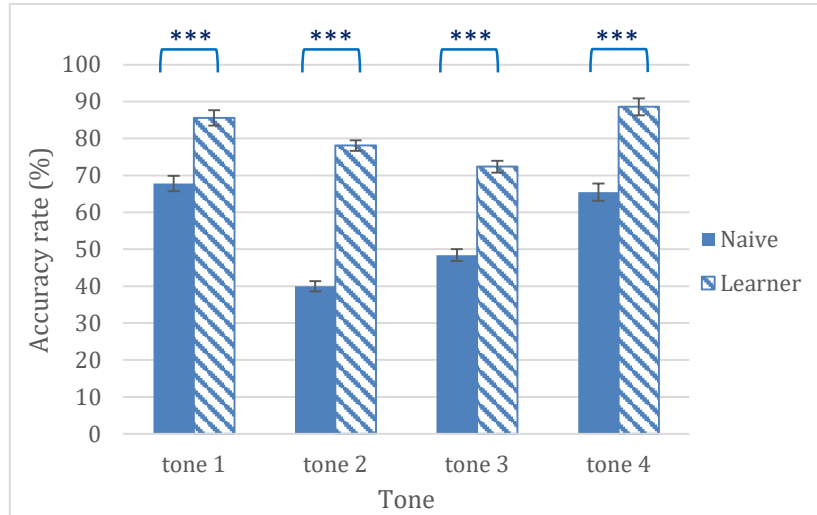


Fig.3. Accuracy rate of each tone in the identification task for the two groups (* means $p < .05$; ** means $p < .01$; *** means $p < .001$).

Regarding the Naive group, there was a significant effect of *tone* [$F(3, 174) = 53.26$, $MSE = 101.90$, $p < .001$] (see Fig. 4).

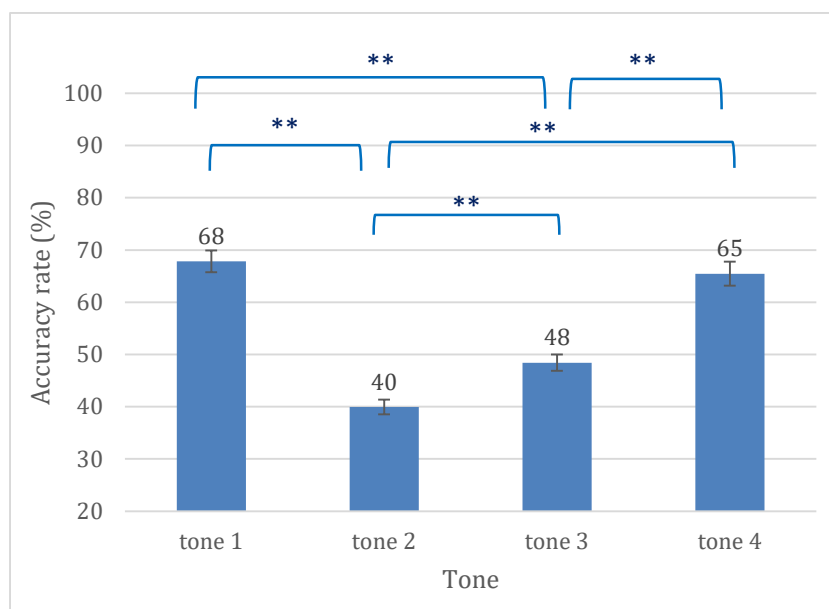


Fig.4. Accuracy rate of each tone in the identification task for the Naive group (* means $p < .05$; ** means $p < .01$; *** means $p < .001$).

From Fig. 4, it showed that the accuracy rate of the Naïve group for Tone 1 was 68%, for tone 2 was 40%, for tone 3 was 48%, and for tone 4 was 65%. Overall, the accuracy rate of Tone 1 as well as Tone 4 was significantly higher than Tone 3, and that of Tone 2 was significantly lower than the other tones.

As for the Learner group, there was also a significant effect of *tone* [$F(3, 174) = 15.87$, $MSE = 101.90$, $p < .001$] (Fig. 5).

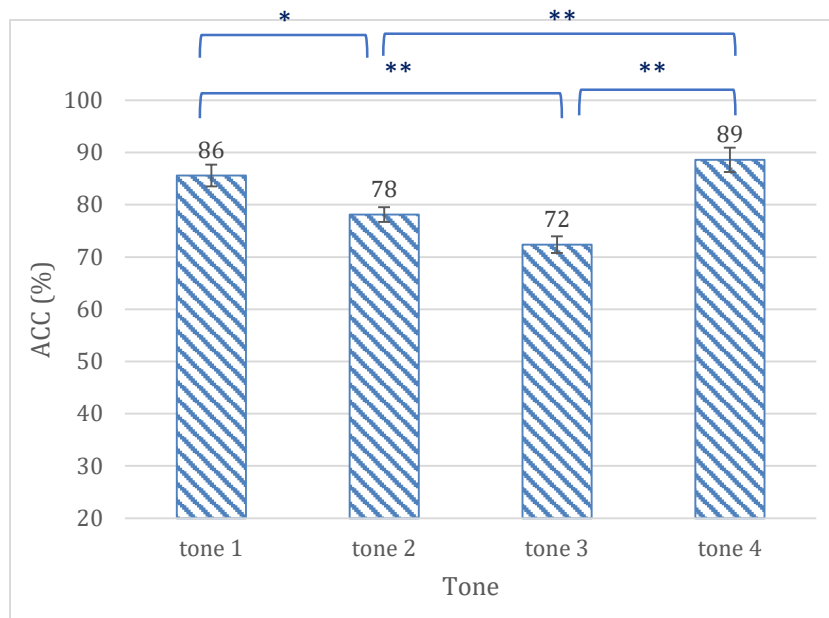


Fig.5. Accuracy rate of each tone in the identification task for the Learner group (* means $p < .05$; ** means $p < .01$; *** means $p < .001$).

In Fig. 5, we see that the accuracy of the Learner group for Tone 1 was 86%, for tone 2 was 78%, for tone 3 was 72%, and for tone 4 was 89%. The overall accuracy rate of Tone 1 as well as Tone 4 was significantly higher than Tone 3 and Tone 2.

Moreover, the mean error rate for each of the 12 tone pairs in the identification task were graphed in Fig.6. (In this study, each tone pair indicated a relationship between a target tone and a response tone. For example, in the T1_T2 pair, Tone 1 is the target tone and Tone 2 is the listeners' response (error). That is, Tone1 is misidentified as Tone 2.

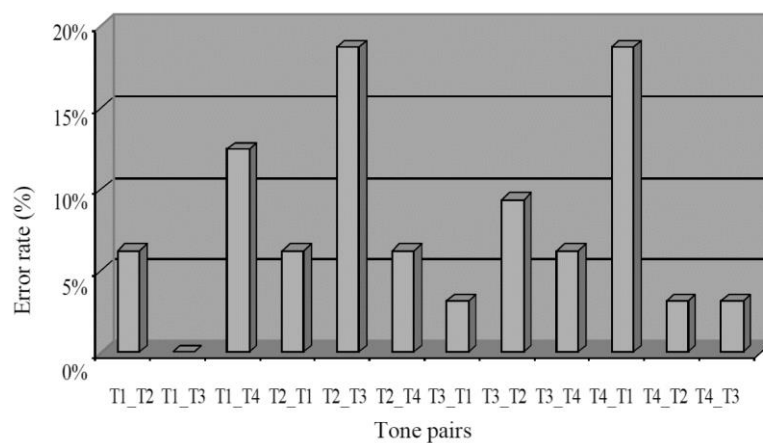


Fig.6. Mean error rate of tonal confusions by the two groups.

It is shown in Fig.6 that errors were mostly found in T2-T3 and T4-T1 pairs, as expected.

The tonal confusions by the naïve listeners and the learners of Mandarin were presented in Fig.7 and Fig.8, respectively.

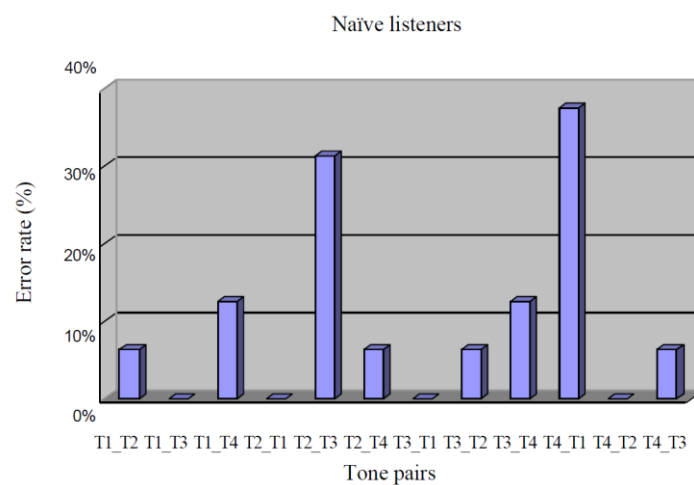


Fig.7. Mean error rate of tonal confusions by the naïve listeners.

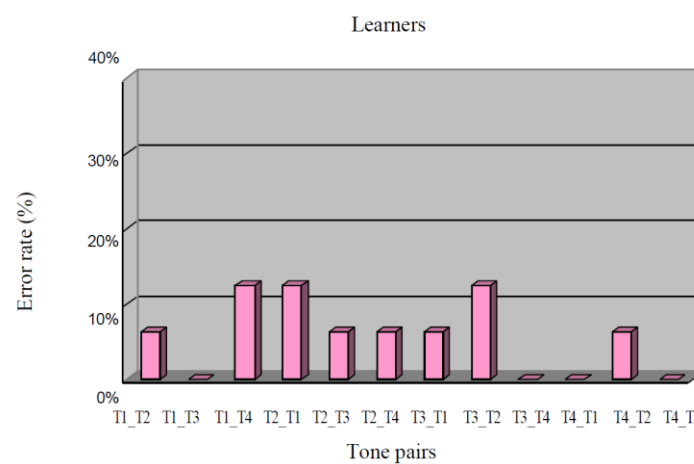


Fig.8. Mean error rate of tonal confusions by the learners.

For the naïve listeners, there were more errors found in T2-T3 (32%) and T4-T1 pairs (38%) than the other pairs; while the learners had more errors in T1-T4 (13%), T2-T1, (13%) and T3-T2 (13%).

3.2 Results of Experiment 2

The similarity rating task was on a 1-5 scale (“1” very different, “5” very similar). The overall results were presented in Fig.9.

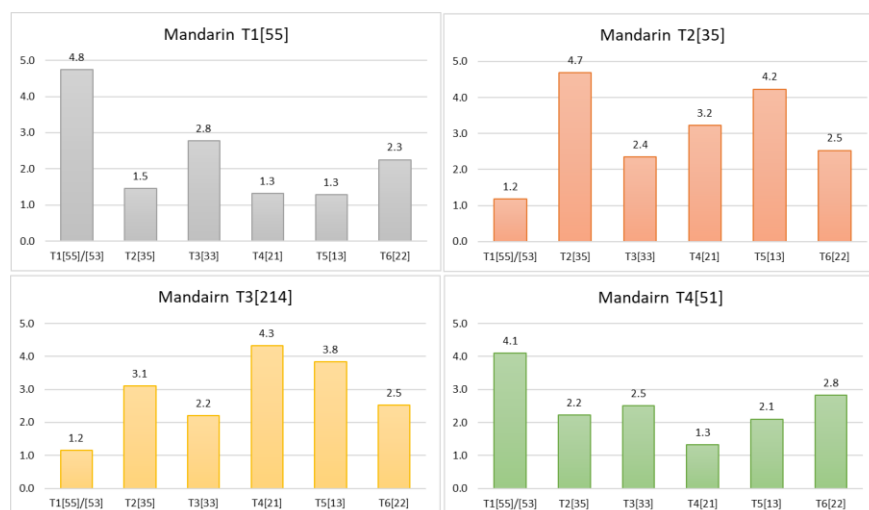


Fig. 9. Average scores of subjective degree-of-similarity ratings between Mandarin and Cantonese tones by Cantonese listener.

From Fig.9, it can be seen that Tone 1 [55] in Mandarin was rated as the most similar to Tone 1 [55/53] in Cantonese. Tone 2 [35] in Mandarin was rated as the most similar to Tone 2 [35], but also similar to Tone 5 [13] in Cantonese. Tone 3 [214] in Mandarin was rated as the most similar to Tone 4 [21], but also similar to Tone 5 [13] in Cantonese. Tone 4 [51] in Mandarin, like Tone 1[55] in Mandarin, was rated as the most similar to Tone 1 [55/53] in Cantonese.

For further analysis, two one-way repeated-measures ANOVAs were conducted for the similarity ratings by the Naïve and Learner groups. The

results showed that for both group, the similarity ratings between individual Mandarin tone and Cantonese tones were significantly different (all $p < .001$), as presented in Table 4.

Table 4. The similarity ratings by the naïve and learner group.

	<i>df</i>	<i>MSE</i>	<i>F</i>	<i>p</i>	η^2	<i>Post hoc analysis</i>
Naïve group (N=30)						
Mandarin T1	5, 145	0.43	131.01	.000	.82	T1 > others (all p values < .001)
Mandarin T2	5, 145	0.58	87.67	.000	.75	T2/T5 > others (all p values < .001)
Mandarin T3	5, 145	0.54	57.31	.000	.66	T4/T5 > others (all p values < .001)
Mandarin T4	5, 145	0.40	75.58	.000	.72	T1 > others (all p values < .001)
Learner group (N=30)						
Mandarin T1	5, 145	0.21	275.25	.000	.91	T1 > others (all p values < .001)
Mandarin T2	5, 145	0.47	106.11	.000	.79	T2 > others (all p values < .01)
Mandarin T3	5, 145	0.34	131.12	.000	.82	T4 > others (all p values < .01)
Mandarin T4	5, 145	0.48	57.76	.000	.67	T1 > others (all p values < .001)

(Note: *post hoc analysis* showed difference between the highest values and others)

In Table 4, we see that for both groups, Mandarin Tone 1 [55] was rated as the most similar to Cantonese Tone 1 [55/53], and Mandarin Tone 4 [55] was also rated as the most similar to Cantonese Tone 1 [55/53]. However, it is shown that the ratings by the two groups were somewhat different for Mandarin Tone 2 and Tone 3. Mandarin Tone 2 [35] was rated as the most similar to both Cantonese Tone 2 [35] and Tone 5 [13] by the Naïve group while that was rated as the most similar to Cantonese Tone 2 [35] by the Learner group. In addition, Mandarin Tone 3 [214] was rated as the most similar to Cantonese Tone 4 [21] and Tone 5 [13] by the Naïve group, but rated as the most similar to Cantonese Tone 4 [21] by the Learner group.

4. DISCUSSION

4.1 Discussion of Experiment 1

As expected, the percent correct for Tone 1 is the highest, i.e., Tone 1 in Mandarin is easier to identify than other tones for Cantonese listeners. Since Tone 1 [55] in Mandarin is categorized as Tone 1 [55] in Cantonese. On the other hand, Tone 2 [35] and Tone 3 [214] have comparatively lower accuracy rates. It is possible that Tone 2 and Tone 3 in Mandarin are uncategorized for Cantonese listeners. In Cantonese, there are two rising tones, Tone 2 and Tone 5. Thus, in terms of pitch contour, the rising Tone 2 [35] in Mandarin is similar to both Tone 2 [35] and Tone 5 [13] in Cantonese. For Tone 3 [214], it falls between Tone 2 [35], Tone 4 [21], and Tone 5 [13] in Cantonese. This result can be accounted for by Flege's (1995) speech learning model (SLM), which proposes that the presence of a native category that is similar but not entirely identical to a non-native one can potentially interfere with the perception of the non-native one. With respect to the tonal confusion patterns, overall results showed more errors for T4-T1 and T2-T3 pairs, as expected. In terms of acoustic properties, Tone 1 [55] and Tone 4 [51] both start at a high level pitch so that Tone 4 is mostly misidentified as Tone 1. Besides, Tone 2 [35] and Tone 3 [214] also have similar acoustic properties since they both have a rising portion of the pitch contours. Li and Thompson (1977) found that L1 children learning Chinese have more problems with Tone 2 and 3 in terms of production. They further mentioned that this confusion can be found in data from perception studies by Kiriloff (1969) and Klatt (1973) as well. Hume and Johnson (2003) measured the perceptual space between the four Mandarin tones and found that Tone 2 and 3 are closer to each other for both Chinese and English speakers. In addition, the tonal confusion patterns reveal the interference of the Cantonese tonal system (L1) on the perception of Mandarin tones. In Cantonese, Tone 1 [55] has a free variant with a falling tone [53]. This is probably why the falling Tone 4 [51] in Mandarin is incorrectly perceived as Mandarin Tone 1 [55] frequently. As for the errors in T2-T3 pair, Cantonese has two rising tones, Tone 2 [35] and Tone 5 [13], and therefore the rising tones in Mandarin, Tone 2 [35] and Tone 3 [214], are easily confused by Cantonese

listeners. The results are consistent with previous findings (Miracle 1989; Shen 1989) that the tone pairs Tone 1-Tone 4 and Tone 2-Tone 3 are the most problematic ones for non-native learners. The direction of confusion for T2-T3 pair is also consistent with Wang, Jongman and Sereno's (2003) study that Tone 2 was incorrectly perceived as Tone 3 more frequently than Tone 3 was incorrectly perceived as Tone 2 by American listeners. However, So (2005) found that in the Tone 2 and Tone 3 pair, Cantonese listeners had a tendency to select Tone 2 as responses most of the time, that is, Tone 3 was mistakenly identified or judged as Tone 2.

It is worth mentioning that the naïve listeners and the learners of Mandarin showed different tonal confusion patterns. The naïve listeners had obviously more errors for T2-T3 and T4-T1 pairs while the learners had more errors for T1-T4, T2-T1, and T3-T2. It is noted that although Tone pair 1 and 4 as well as Tone pair 2 and 3 are the most confusing pairs for both the naïve listeners and learners, the direction of confusion is different for these two groups. The naïve listeners misidentified Tone 2 as Tone 3, Tone 4 as Tone 1 frequently, but not so much vice versa. The learners, however, misidentified Tone 3 as Tone 2, Tone 1 as Tone 4 more frequently than the reverse. For the tone pair 2 and 3, the naïve listeners showed the same pattern (misidentified Tone 2 as Tone 3) as the American listeners did in the study by Wang et al. (2003), while the learners showed the same pattern as the Cantonese listeners did in So's (2005) study, which claims that Cantonese listeners had a tendency to identify Tone 3 as Tone 2.

4.2 Discussion of Experiment 2

The data from the similarity rating task demonstrates how similar each Mandarin tone is to a Cantonese tone. The results show that Tone 1 and Tone 4 in Mandarin were both rated as the most similar to Tone 1 in Cantonese. The two tones in Mandarin are similar to a single tone in Cantonese and therefore in this case it was seen as Single Category (SC) assimilation, which describes that both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but as equally good exemplars of that category (Best & Tyler 2007). According to Best's (1995) PAM, single category L2 contrast assimilation hinders

listeners' perception of non-native sounds. The results from the similarity ratings further explain why the tonal confusions were mostly found in Tone 1 and Tone 4 pairs, and therefore lend support to PAM.

The results also show that Tone 2 and Tone 3 in Mandarin were not clearly rated as a single tone and were rated as similar to more than one single tone in Cantonese. Although Mandarin Tone 2 was rated as the most similar to Tone 2 in Cantonese, the average scores for Cantonese Tone 2 and Tone 5 by the naïve group were not significantly different. The result indicated that Mandarin Tone 2 were perceived similar to both Tone 2 and Tone 5 in Cantonese by the naïve listeners. The situation was also found for Tone 3. Mandarin Tone 3 was rated as the most similar to Tone 4 in Cantonese, but the average score for Cantonese Tone 4 and Tone 5 by the naïve group were not significantly different. The result infers that Tone 3 in Mandarin was perceived similar to both Tone 4 and Tone 5 in Cantonese by the naïve listeners. Thus, Mandarin Tone 2 and Tone 3 were perceived as having a mixture of more modest similarities to several L1 phonological categories (Best & Tyler 2007). According to PAM, if the uncategorized L2 phones have similarities to the same set of L1 phonemes, then the listeners should find them difficult to discriminate and perceive relevant L2 lexical-functional differences (Best & Tyler 2007). The results for Mandarin Tone 2 and Tone 3 provide the empirical data to examine this claim. Tone 2 and Tone 3 in Mandarin were both perceived as similar to the same set of Cantonese tones, Tone 2 and Tone 5 as well as Tone 4 and Tone 5, and thus they were close to each other in phonological space with cross-language contrast. Moreover, Hume and Johnson (2003) found that Tone 2 and 3 are closer to each other in the perceptual space between the four Mandarin tones for both Chinese and English speakers. Recalling the results of identification from Experiment 1, percent correct for Tone 2 and Tone 3 were lower than the others. It is not only because the acoustic properties of the two tones are pretty similar in nature, but also they are confused with the native tonal system for Cantonese listeners. The results here provide further support to PAM on the suprasegmental level.

4.3 General Discussion

Regarding the hypotheses proposed earlier, the Cantonese listeners did better in identifying Mandarin Tone 1 [55]. It is probably because Tone 1 [55] in Mandarin was categorized as Tone 1 [55/53] in Cantonese. The two tones are similar not only in their acoustic properties but also in perception by Cantonese listeners. Based on the results of the similarity rating task, it was observed that Mandarin Tone 1 exhibited a high degree of perceptual assimilation to Cantonese Tone 1. It was also found that the accuracy rate for Tone 1 [55] was also higher than the other tones. Tone 4 [51] in Mandarin seems to be also categorized as Tone 1 [55/53] in Cantonese since Cantonese Tone 1 [55] has a free variant with falling tone [53], which is similar to Mandarin Tone 4 [51], and Mandarin Tone 4 is also rated as very similar to Tone 1 in Cantonese.

As for the tonal confusion patterns, the overall results were consistent with previous findings (So 2005) that tonal errors were mostly found in T1-T4 and T2-T3 pairs (bidirectional) though more errors in T4-T1 and T2-T3 pairs were found in this study. Further examining the data, it was found that the naïve listeners and the learners showed different tonal confusion patterns in the direction of confusion. The naïve listeners had more errors in T4-T1 and T2-T3 pairs, but the learners had more errors in T1-T4 as well as T3-T2 pairs and even no error was found in T4-T1. In terms of acoustic properties, Tone 1 and Tone 4 both start as a high level pitch so that they can be confusing with each other. Moreover, the tonal confusions of Tone 1 and Tone 4 reveal L1 interference in that Tone 1 and Tone 4 in Mandarin are both perceived as similar to Tone 1 with free variants [55/53] in Cantonese. This case can be explained by PAM. According to PAM (Best 1995), when two non-native phones assimilate to a single native phone, described as Single Category (SC) assimilation, it is difficult for listeners' to correctly perceive the non-native phones.

Concerning the tonal confusions of Tone 2 and Tone 3, more errors in this tone pair were also found for the listeners with other language background. Chen (1997) found that English-speaking learners often confused Tone 2 and Tone 3 in perception. Even for native speakers of Mandarin, Tone 2 and Tone 3 were indeed perceptually more confusable than any other tone pairs (Huang 2007). This contrast between Tone 2 and

Tone 3 may be intrinsically more difficult. Tone 2 and Tone 3 have similar acoustic properties with a rising portion of pitch contours, as results in easy confusions between the two tones. In addition, comparing the tonal similarities between Mandarin and Cantonese in perception, Tone 2 and Tone 3 in Mandarin were both perceived as similar as more than one single tone in Cantonese. Mandarin Tone 2 [35] was rated as similar to Cantonese Tone 2 [35] and Tone 5 [13]; Mandarin Tone 3 [214] was rated as similar to Cantonese Tone 4 [21] and Tone 5 [13]. Thus, the two tones were perceived as belonging to the same set of L1 tonemes for Cantonese listeners. Just as PAM predicts, the two tones are difficult to discriminate and to perceive relevant L2 lexical differences.

5. CONCLUDING REMARKS

The findings in this study suggest that the tonal errors may result from not only the similar acoustic properties of the tone pairs but also perceptual assimilation between L1 and L2 tonal contrasts. Therefore, the results should lend support to PAM at the suprasegmental level.

The data in this study were limited and some aspects need further investigation. For future studies, the participants should include native listeners of Mandarin as a control group. More participants with different backgrounds, such as L2 proficiency and musical training, for learners should be recruited. In addition, the other direction of perception, i.e., Cantonese tone perception by Mandarin listeners, is worth conducting. Also, future research should include more than one exemplar for each tone since there could be frequency effects or some other confounding variables at play.

To further examine the effects of prosodic similarities on L2 perception in terms of the suprasegmental level, a cross-linguistic study may also include non-native listeners of other tone languages, e.g., Thai, Vietnamese, and non-tone languages, e.g., English, French, as well as pitch-accent languages, e.g., Japanese, Korean. Then, the possibility of applying PAM to suprasegmental level and L2 prosodic perception will be better studied.

REFERENCES

- Bauer, Robert, and Paul Benedict. 1997. *Modern Cantonese Phonology*. Berlin and New York: Mouton de Gruyter.
- Bent, T. 2005. *The perception and production of non-native prosodic categories*. Ph.D. Dissertation, Department of Linguistics, Northwestern University, Evanston, IL.
- Best, Catherine. 1995. A direct realist view of cross-language speech perception. *Speech perception and Linguistic Experience: Issues in cross-language research*, ed. by Winifred Strange, 171-204. New York.
- Best, Catherine, and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross language perception of approximants. *Journal of Phonetics* 20:305-331.
- Best, Catherine, and Michael Tyler. 2007. Nonnative and second-language speech perception. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, ed. by Ocke-Schwen Bohn, and Murray J. Munro, 13-34. Amsterdam: John Benjamins.
- Chandrasekaran, B., Sampath, P.D., & Wong P.C.M. 2010. Individual variability in cue-weighting and lexical tone learning. *Journal of Acoustical Society of America*, 128: 456-465.
- Chao, Yuan-ren. 1930. A System of Tone Letters. *La Maître phonétique* 45:24-27.
- Chen, M. Y. 2000. *Tone Sandhi: Patterns Across Chinese Dialects*. Cambridge: Cambridge University Press.
- Chen, Qinghai. 1997. Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association* 32: 21-39.
- Cutler, Anne, and Takashi Otake. 1994. Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language* 33:824-844.
- Ding, H. 2012. Perception and production of Mandarin disyllabic tones by German learners. In *Proceedings of 2012 Speech Prosody*, Shanghai, pp. 378-381.
- Gandour, J. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149-174.
- Gandour, J. 1984. Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics* 12: 235-261.
- Flege, James E. 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89(1):395-411.
- Flege, James E. 1993. Production and perception of a novel second-language phonetic contrast. *Journal of the Acoustical Society of America* 93(3):1589-1608.
- Flege, James E. 1995. Second language speech learning: Theory, findings and problems. *Speech perception and linguistic experience: Theoretical and methodological issues*, ed. by Winifred Strange, 233-277. Baltimore: York Press.
- Flege, James E., Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25: 437-470.

- Flege, James E., Naoyuki Takagi, and Virginia Mann. 1995. Japanese adults can learn to produce English /O/ and /I/ and accurately. *Language and Speech*, 38: 25-55.
- Fon, Janice, and Wen-yu Chiang. 1999. What does Chao have to say about tones. *Journal of Chinese Linguistics* 27(1): 13-37.
- Francis, Alexander L., Valter Ciocca, Lian Ma, and Kimberly M. Fenn. 2008. Perceptual learning of Cantonese lexical tones by tonal and non-tonal language speakers. *Journal of Phonetics* 36: 268-294.
- Gandour, Jackson. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149-174.
- Gandour, J. 1984. Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics* 12: 235-261.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. 2000. A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12(1): 207-222.
- Gandour, J., Wong, D., & Hutchins, G. 1998. Pitch processing in the human brain is influenced by language experience. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 9(9): 2115-2119.
- Guion, S.G. Pederson E. 2007. Investigating the role of attention in phonetic learning. In O.-S. Bohn, M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege*, pp. 57-77. John Benjamins, Amsterdam.
- Halle, Pierreb, Yueh-chin Chang, and Catherine Best. 2004. Identification and discrimination of Mandarin Chinese tones vs. French listeners. *Journal of Phonetics* 32: 395-421.
- Hao, Y. C. 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40: 269-279.
- Huang, Tsan. 2007. Perception of Mandarin tones by Chinese-and English-speaking listeners. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 1797- 00.
- Hume, Elizabeth, and Keith Johnson. 2003. The impact of partial phonological contrast on speech perception. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pp. 2385-88.
- Jongman, A., Wang, Y., Moore, C., & Sereno, J.A. 2006. Perception and production of Mandarin Chinese tones. In P. Li, E. Bates, L.H. Tan, & O. Tseng (Eds.), *The Handbook of East Asian Psycholinguistics*, pp.209-217. Cambridge University Press.
- Jun, Sun-Ah, and Mira Oh. 2000. Acquisition of second language intonation. In *ICSLP- 2000*(4), pp.73-76.
- Kiriloff, C. 1969. On the auditory perception of tones in Mandarin. *Phonetica*, 20, 2-4: 63-67.
- Klatt, D. (1973). Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. *Journal of the Acoustical Society of America*, 53: 8-16.
- Kuhl, Patricia K. 2000. A new view of language acquisition. In *Proceedings of the National Academy of Science* 97, pp.11850-11857

- Lee, Yuh-shiow, Douglas A. Vakoch, and Lee H Wurm. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research* 25(5): 527-544.
- Li, Charles N., and Sandra A. Thompson. 1977. The acquisition of tone in Mandarin- speaking children. *Journal of Child Language* 4:185-99.
- McMurray, B., Tanenhaus, M. K., & Aslin, R. N. 2002. Gradient effects of within-category phonetic variation on lexical access. *Cognition*, 86: B33-B42.
- McMurray, B., Tanenhaus, M. K., & Aslin, R. N. 2009. Within-category VOT affects recovery from “lexical” garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, 60: 65-91.
- Miracle, Charles. 1989. Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association* 24:49-65.
- Mok, P. P. K. & Wong, P. W. Y. 2010. Perception of the merging tones in Hong Kong Cantonese: preliminary data on monosyllables”, *Speech Prosody 2010*, 100916, pp.1-4.
- Munro, Murray J., and Ocke-Schwen Bohn. 2007. The study of second language speech: A brief overview. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, ed. by Ocke-Schwen Bohn, and Murray J. Munro, pp. 3-11. Amsterdam: John Benjamins.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Pallier, Christophe, Laura Bosch, and Nuria Sebastian-Gallés. 1997. A limit on behavioral plasticity in speech perception. *Cognition* 64: B9-B17.
- Qin, Z., Tremblay, A. & Zhang, J. 2019. Influence of within-category tonal information in the recognition of Mandarin-Chinese words by native and non-native listeners: An eye-tracking study. *Journal of Phonetics*, 73: 144-157.
- Shen, Xiao-nan Susan. 1989. Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association* 24:27-47.
- Shen, J., Deutsch, D., & Rayner K. 2013. On-line perception of Mandarin Tones 2 and 3: Evidence from eye movements. *Journal of the Acoustical Society of America*, 133, 3016-3029.
- So, C. & Best, C. 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and speech*, 53: 273-93.
- So, Connie K. 2005. The influence of L1 prosodic background on the learning of Mandarin tones: Patterns of tonal confusion by Cantonese and Japanese naïve listeners. In *Proceedings of the 2005 annual conference of the Canadian Linguistic Association*.
- Stagray, James R., and David Downs. 1993. Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics* 21:143-63.
- Strange, Winifred. 2007. Cross-language phonetic similarity of vowels. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, ed. by Ocke-Schwen Bohn, and Murray J. Munro, pp.35-55. Amsterdam: John Benjamins.

- Tsukada, Kimiko. 2005. Cross-language speech perception of final stops by Australian- English, Japanese and Thai listeners. In *Proceedings of ISCA Workshop Plasticity Speech Perception*, London, pp.244-247.
- Tu, J., Hsiung, Y., Cha, J., Wu, M. & Sung, Y. 2016. Tone production of Mandarin disyllabic words by Korean learners. In *Proceedings of Speech Prosody 8*, pp.375-379.
- Tu, J., Hsiung, Y., Wu, M. & Sung, Y. 2014. Error patterns of Mandarin disyllabic tone by Japanese learners. In *Proceedings of Interspeech-2014*, pp.2558-2562.
- Vince, T. 1984. An experimental investigation of tone and intonation in Cantonese. *Phonetica*, 33: 368-392.
- Walley, Amanda C., and James E. Flege. 1998. Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics* 27:307-332.
- Wang, Yue, Michelle M. Spence, Allard Jongman, and Joan A. Sereno. 1999. Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America* 106: 3649-3658.
- Wang, Yue, Allard Jongman, and Joan A. Sereno. 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America* 113:1033-1043.

Jung-yueh Tu
 PhD/MA Program in Teaching Chinese as a Second Language
 National Chengchi University
 Taipei, Taiwan 116, ROC
 jytu@nccu.edu.tw

Appendix A – List of stimuli in Experiment 1 Mandarin words

pinyin	IPA	T1 [55]	T2 [35]	T3 [214]	T4 [51]
bi	[bi]	逼	鼻	比	必
ling	[liŋ]	拎	靈	領	令
wei	[weɪ]	威	圍	偉	味
fēn	[fən]	分	焚	粉	份

Appendix B – List of stimuli in Experiment 2

Mandarin	[fən]	T1 [55]	T2 [35]	T3 [214]	T4 [51]		
		分	焚	粉	份		
Cantonese	[fən]	T1 [55/53]	T2 [35]	T3 [33]	T4 [21]	T5 [13]	T6 [22]
		分	粉	訓	焚	奮	份

Mandarin	[fu]	T1 [55]	T2 [35]	T3 [214]	T4 [51]		
		夫	福	府	婦		
Cantonese	[fu]	T1 [55/53]	T2 [35]	T3 [33]	T4 [21]	T5 [13]	T6 [22]
		夫	苦	富	符	婦	負

二語聲調處理歷程的感知同化效應

杜容珣
國立政治大學

本研究探討粵語母語者對華語聲調的感知機制。實驗一，粵語母語者聽到華語聲調後辨識所聽到的是哪個聲調。實驗二，評比所聽到的華語和粵語聲調配對的相似度。研究發現，聲調感知的混淆偏誤可能不只來自於聲調之間相近的聲學特徵，也來自於母語和二語聲調對比之間的感知同化程度。基於研究結果，本文討論母語韻律系統對於二語感知的作用，以及感知同化類型如何預測字詞聲調範疇的感知表現。

關鍵字:感知同化、華語聲調、聲調處理歷程、二語感知