PERCEPTUAL ASSIMILATION MODEL FOR SUPRASEGMENTALS PREDICTS CANTONESE TONE SEQUENCE RECALL IN PUNJABI MINORITIES

Doris Yu[1], Veronica Ka Wai Lai[2], William Choi[1]

The University of Hong Kong[1], University of Manitoba[2]
dorisyuu@connect.hku.hk, veronica.lai@umanitoba.ca, willchoi@hku.hk

ABSTRACT

Research has suggested that the Perceptual Assimilation Model for Suprasegmentals (PAM-S) predicts non-native tone discrimination. This study investigates whether PAM-S applies to sequence recall. Twenty-two Punjabi speakers completed a Cantonese tone perceptual assimilation task and a Cantonese tone sequence recall task. In the perceptual assimilation task, Punjabi speakers assimilated the Cantonese level tone contrast (T1-T3) and the contour tone contrast (T4-T5) as single category (SC) and uncategorized-categorized pair without overlap category (UC_{no}), respectively. In the sequence recall task, they recalled T4-T5 (UC_{no}) more accurately than T1-T3 (SC) across different sequence lengths. The findings are consistent with the prediction of PAM-S on sequence recall. Moreover, memory load does not seem to limit its applicability.

Keywords: The Perceptual assimilation model for suprasegmentals, sequence recall, Cantonese tone.

1. INTRODUCTION

About 8% of the population in Hong Kong are non-Chinese ethnic minorities with diverse language backgrounds (e.g., Hindi, Punjabi, and Urdu) [1]. Research has found that non-Chinese ethnic minorities have difficulties learning Cantonese [2-4], particularly in Cantonese tone perception [5]. With the increase in learning needs in Cantonese as a second language [6], however, the perceptual patterns across groups with various native language backgrounds have not been fully explored. Thus, there is a dearth of research examining the perceptual characteristics of various non-Chinese ethnic minorities who learn Cantonese as their second language (L2). Among the native languages non-Chinese ethnic minorities speak, Punjabi is a tone language whereas Urdu and Hindi are not. Additionally, the question of how the native tone system of Punjabi speakers influences their phonological perception has yet to be explored [7]. From the theoretical perspective, previous studies have not examined whether PAM-S applies to the tone sequence recall task which taps into phonological perception [8-12]. Hence, we examined how Punjabi listeners assimilate Cantonese tones and recall Cantonese tone sequences.

PAM-S predicts the discrimination patterns of non-native suprasegmentals based on the assimilation patterns [13, 14]. For perceptual assimilation, listeners' native prosodic systems may or may not categorize non-native suprasegmentals. The assimilation patterns can be classified into five types: assimilation of two non-native categories into two native categories (two-category assimilation, TC), assimilation of two non-native categories into one native category (single-category assimilation, SC), assimilation of two non-native categories into one native category with different assimilation goodness (category goodness difference in assimilation, CG), assimilation of one categorized non-native category and one uncategorized non-native category (uncategorized-categorized, UC), and assimilation of two uncategorized non-native categories (uncategorized-uncategorized, UU) [13, 14].

Regarding the discriminability of the assimilation pairs, PAM-S predicts that non-native speakers most easily discriminate TC; easily discriminate UC with no native category overlap (UC_{no}) and UU with no native category overlap (UU_{no}); and fairly easily discriminate CG [13, 14]. Additionally, non-native speakers poorly discriminate SC, UC with native category overlap (UC_{o}), UU with partial native category overlap (UU_{o}), and UU with the same category overlap (UU_{s}) [13, 14].

Previous studies put forward that the AX tone discrimination task cannot probe into phonological processing even with memory load manipulations [15, 16]. In the AX tone discrimination task, listeners have to store the first stimulus in short-term memory, and use phonetic details to judge whether the second stimulus is the same as the first stimulus [15, 16]. In previous studies, listeners' discrimination accuracy was not influenced by short (500 ms) or long (2000 ms) interstimulus intervals [15, 16]. These findings suggest that AX tone discrimination task under short and long memory loads can uphold the PAM-S at the phonetic level but not at the phonological level [15,16]. In this regard, the tone discrimination task

under memory load manipulations limits the applicability of PAM-S in phonological perception. To extend the applicability of PAM-S at the phonological level, it is necessary to adopt a more robust test of the phonological mode of perception.

Sequence recall task is a potent tool to probe into the phonological mode of perception. In the familiarization session, listeners temporarily encode nonnative speech sounds into short-term phonological representations [8, 12]. In the test session, listeners have to use the phonological representations to efficiently encode and retain the information in short-term memory for recalling the sequence. In addition, using phonetic variability in sequence recall task can prevent participants from using acoustic features, since stimuli produced by different talkers pose more difficulties for listeners to employ acoustic features and listeners thus have to use phonological representations to efficiently encode the information [8-12]. Therefore, the sequence recall task with phonetic variability is a robust short-term memory test for phonological perception.

Cantonese and Punjabi are both tone languages. Punjabi comprises three lexical tones: high falling tone, level tone, and low rising tone [17-19]. Likewise, Cantonese consists of six lexical tones: high level tone (T1), high rising tone (T2), mid-level tone (T3), low falling tone (T4), low rising tone (T5), and low level tone (T6) [20]. Based on the similarities and distinctions between Cantonese and Punjabi lexical tone systems, we formulated a hypothesis on how Punjabi listeners would assimilate Cantonese tones into the Punjabi tone system (see Table 1). To avoid prolonged testing and exhaustion effects, we only employed T1-T3 and T4-T5 pairs in the tone sequence recall task. Since T1, T3, and the Punjabi level tone are level tones, we expected that T1 and T3 would be assimilated as the Punjabi level tone. As such, the T1-T3 contrast is an SC pair. On the other hand, T4 and T5 are contour tones with different pitch contours, i.e., low falling contour and low rising pitch contour, respectively. As the tonal patterns of T5 and the Punjabi low rising tone are similar, we assumed that T5 would be assimilated as the Punjabi low rising tone. Given that no low falling tonal pattern was found in the Punjabi tone system, we expected that T4 would not be assimilated as a single Punjabi tone, but rather it would be equally assimilated as the Punjabi level tone and high falling tone because T4 has a similar tone onset to the Punjabi level tone and a similar tone contour to the high falling tone. For this reason, we hypothesized that T4 and T5 were assimilated to a UC_{no} pair. In short, we suggested that the T1-T3 pair would represent SC pair and T4-T5 would represent as UC_{no} pair (see Table 2), and both tone pairs would yield empirically testable predictions of PAM-S.

If PAM-S fits well for phonological perception in the sequence recall task, the recall should be better for UC_{no} than for SC. Otherwise, Punjabi listeners might not recall UC_{no} better than SC at the phonological level in the sequence recall task.

Tone	Punjabi low-rising	Punjabi level	Punjabi high-falling	Uncate-gorized	None
T1		+			
T2	+				
T3		++			
T4			+		
T5	++				
T6		+			

Table 1: Hypothesis of Cantonese tone assimilation patterns[1]

Tone	T1	T2	T3	T4	T5
T2	TC				
T3	**SC**	TC			
T4	UC_o	UC_{no}	UC_o		
T5	TC	SC	TC	**UC_{no}**	
T6	CG	TC	SC	UC_o	TC

Table 2: Hypothesis of assimilation pairs

Altogether, this study set out to examine whether PAM-S applies to tone sequence recall. In particular, we investigated (i) how Punjabi listeners assimilate the six Cantonese tones into their native tone system and (ii) whether Punjabi listeners perform better on recalling the UC_{no} pair than the SC pair.

2. METHODS

2.1. Subjects

Twenty-two Punjabi speakers (M_{age} = 22.5 years old, SD = 11.7 years) living in Hong Kong participated in the experiment. All subjects were native Punjabi speakers and grew up in Punjabi-speaking households. They identified Cantonese as their second language. They self-reported their Punjabi (M = 4.6, SD = 0.6) and Cantonese proficiencies (M = 2.3, SD = 0.1) on a 5-point Likert scale (1 = naïve, 5 = native), with Punjabi proficiency significantly higher than that of Cantonese ($t(21) = 9.63$, $p < .001$). Punjabi (M = 4.2, SD = 1.1) was their dominant language for communication while Cantonese was seldom used (M = 2.6, SD = 1.1) on a Likert scale (1 = never, 5 = always), and the difference between them attained statistical significance ($t(21) = 4.9$, $p < .001$). Their average length of residency in Hong Kong was 14.3 years (SD = 7.4 years). All subjects reported having normal hearing.

2.2. Cantonese Tone Assimilation Task

Subjects were visually and audibly presented with three Punjabi words (/tʃá/ *tea*, /tʃã/ *enthusiasm*, and /tʃà/ *peep*) with their corresponding tone symbols. To ensure all subjects were familiar with the three Punjabi tone symbols, they had to successfully match the Punjabi words to the tone symbols before starting the test.

The stimuli were 12 Cantonese words (/fu-T1/, /fu-T2/, /fu-T3/, /fu-T4/, /fu-T5 /, /fu-T6 /, /ji-T1/, /ji-T2/, /ji-T3/, /ji-T4/, /ji-T5/, /ji-T6/) naturally produced by two native Cantonese female adults and two native Cantonese male adults. On each trial of the test phase, a Cantonese word was played once with four options (the three Punjabi tone symbols and a "none" option) presented on the screen. Each subject was required to select a sign and indicate the degree of the heard tone that matched the selected option on a seven-point scale (1 = poor, 7 = perfect). The test phase contained 72 trials (12 Cantonese words × 6 repetitions). The reliability of the assimilation task was satisfactory (Cronbach's α = .8).

2.3. Cantonese Tone Sequence Recall Task

The sequence recall task was adopted from [10], with a new set of stimuli. There were two pairs (/ji-T4/ and /ji-T5/, /ji-T1/, and /ji-T3/) with five sequence lengths in this task. Based on the results of the assimilation task, these two pairs could be defined as a UC_{no} pair (/ji-T4/ and /ji-T5/) and an SC pair (/ji-T1/ and /ji-T3/). The stimuli were naturally produced by the same Cantonese-speaking adults as aforesaid. In the familiarization phase, subjects were able to associate the /ji-T1/ and /ji-T5/ associated with key 1, and the /ji-T3/ and /ji-T4/ with key 2.

On each trial, a sequence of two to six Cantonese words (e.g., /ji-T1/, /ji-T3/, /ji-T1/, /ji-T3/) with 80ms inter-stimulus interval was played once. To reduce subjects' reliance on acoustic memory, each sequence was followed by a pure tone with an interval of 80 ms. The order of genders was counterbalanced across tones in each trial. Each subject was required to replicate the sequence of tones by pressing the associated keys (e.g., 1212 for /ji-T1/, /ji-T3/, /ji-T1/, /ji-T3/). One point was given only when the consecutive answer correctly matched the sequence. Feedback was provided only in the practice phase. The test phase included 60 trials (2 assimilation pairs × 5 sequence lengths × 6 trials per sequence length). The reliability of the sequence recall task was high in the UC_{no} pair (α = .90) and the SC pair (α = .87). To eliminate the impact of a carryover effect, the Cantonese tone sequence recall was conducted before the assimilation task.

3. RESULTS

3.1. Cantonese Tone Assimilation Task

We adopt the categorization criteria set by [13, 14]. If a Cantonese tone matches one of the assimilation options with a percentage higher than chance level (25%), then the Cantonese tone is considered to be assimilated as a Punjabi tone or none category [13, 14]. Additionally, the mean of category goodness should also be higher than five for the sake of stringency.

To explore the assimilation patterns of each Cantonese tone by Punjabi speakers, one-way ANOVAs were utilized on each Cantonese tone with assimilation option as the within-subject factor. For T3, Greenhouse-Geisser adjustment was used to correct for a sphericity violation. The Shapiro-Wilk test indicated that five Cantonese tones did not follow normally distributed patterns. Since one-way ANOVA is robust to deviations from normal distribution [21], we presented the results obtained from it. The results indicated that the main effects of the assimilation options were significant for each Cantonese tone: T1, $F(3, 33) = 19.56$, $p < .001$, $\eta_p^2 = .64$; T2, $F(3, 33) = 15.89$, $p < .001$, $\eta_p^2 = .59$; T3, $F(1.42, 15.62) = 20.93$, $p < .001$, $\eta_p^2 = .66$; T4, $F(3, 33) = 26.45$, $p < .001$, $\eta_p^2 = .71$; T5, $F(3, 33) = 11.71$, $p < .001$, $\eta_p^2 = .52$; T6, $F(3, 33) = 35.15$, $p < .001$, $\eta_p^2 = .76$. Levene's test revealed that the homogeneity of variance assumption was not met for two Cantonese tones. In this respect, we used one-way Welch's ANOVA, which yielded consistent results.

It is evident from the results that T1, T3, and T6 had similar assimilation patterns. The three tones were assimilated more frequently as the Punjabi level tone, $ps < .01$, than to the Punjabi high falling tone, the Punjabi low rising tone, and none category. T2 and T5 also showed similar assimilation patterns. Both tones were assimilated more frequently as the Punjabi low rising tone, $ps < .05$, compared with the Punjabi level tone, the Punjabi high falling tone, and none category. T4 and T5 had different assimilation patterns. T4 was assimilated more frequently as the Punjabi high falling and the level tone, $ps < .01$, than as the Punjabi low rising tone and none category. Unlike T4, T5 was assimilated more frequently as the Punjabi low rising tone, $ps < .05$, than to the Punjabi level tone, the Punjabi high falling tone, and none category.

The results were mainly consistent with our assimilation hypothesis. Of interest to our study is that, the T1-T3 pair is assimilated as SC, and the T4-T5 pair is assimilated as UC_{no}. Figure 1 shows the assimilation percentages of four Cantonese tones, and

Table 3 indicates the assimilation patterns obtained from pairwise comparisons.

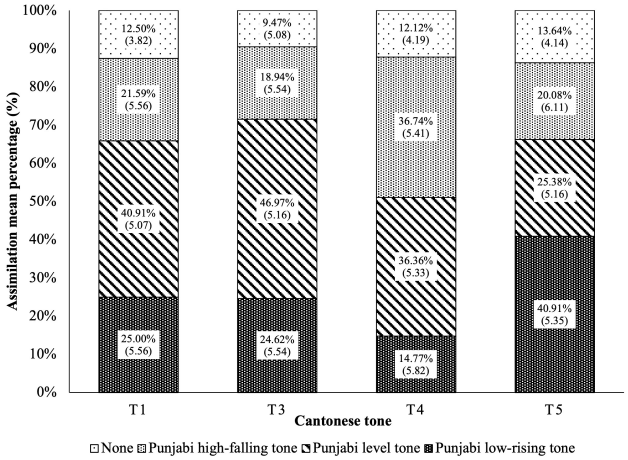

Figure 1: Assimilation percentages of Cantonese T1, T3, T4, and T5 and their means of category-goodness in parentheses

Tone	Punjabi low-rising	Punjabi level	Punjabi high-falling	Uncate-gorized	None
T1		++			
T3		++			
T4		+	+	++	
T5	++				

Table 3: Assimilation patterns of Cantonese T1, T3, T4, and T5

3.2. Cantonese Tone Sequence Recall Task

A two-way ANOVA was conducted on mean accuracies with assimilation pairs (UC_{no} and SC) and sequence length (two, three, four, five, six) being the within-subject factors. As Levene's test showed a violation of the assumption of homogeneity of variance, we performed a square root transformation on the mean accuracy to eliminate the violation. Then, ANOVA was done on the transformed percentage accuracy. The Shapiro-Wilk test of normality showed that the transformed accuracy was normally distributed. Mauchly's test indicated that the assumption of sphericity has been met, $\chi^2 = 6.65$, $df = 9$, $p = 0.844$. The results showed a significant main effect of assimilation pair, $F(1, 21) = 48.84$, $p < .001$, $\eta_p^2 = .70$. Post-hoc Bonferroni tests showed the mean percentage of UC_{no} pair was significantly higher than the SC pair ($p < .001$). We also found the significant main effect of sequence length, $F(4, 84) = 35.21$, $p < .001$, $\eta_p^2 = .63$. Specifically, pairwise comparisons with Bonferroni adjustments showed that the mean accuracy of two-tone sequence length was significantly higher than the three-, four-, five-, and six-tone sequence lengths ($ps < .05$), and three-tone sequence length was significantly higher than the four-, five-, and six-tone sequence lengths ($ps < .01$), but there was no significant difference in four-tone sequence length compared to the five- and six-tone sequence lengths ($ps \geq .074$). The interaction between assimilation pair and sequence length was not statistically significant, $p = .656$. Figure 2 shows the mean accuracy (not transformed) of two assimilation pairs in the sequence recall task.

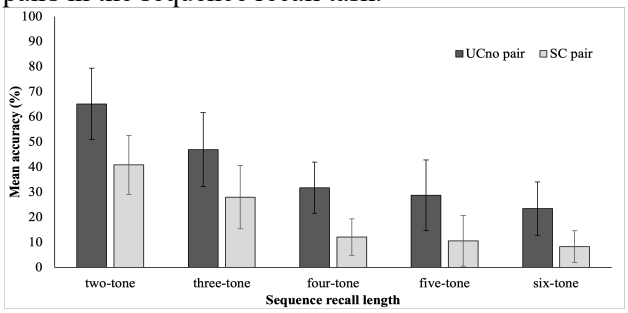

Figure 2: Mean accuracy of the UCno and SC tone pairs in the sequence recall task.

Additionally, we conducted supplementary analyses with outlier removal of the length of residency and found the results were consistent with the assimilation patterns and sequence recall performance.

4. DISCUSSION

The sequence recall performance was consistent with PAM-S' prediction on tone discrimination. In PAM-S, the discriminability of the UC_{no} pair is higher than the SC pair at low and high perceptual levels [13-15]. In this study, Punjabi listeners assimilated Cantonese T1-T3 as an SC pair and T4-T5 as a UC_{no} pair. In the sequence recall task, they exhibited a better performance in recalling T4-T5 (UC_{no}) than T1-T3 (SC), aligning with the prediction of PAM-S. The present study extends the predictive potential of PAM-S for predicting SC and UC_{no} pairs in L2 tone sequence recall. In particular, memory load does not limit the applicability of PAM-S in the tone sequence recall task.

5. CONCLUSION

To sum up, our findings reflect the applicability of PAM-S on tone sequence recall. Particularly, Punjabi speakers recalled the UC_{no} pair significantly better than the SC pair. Given that we only included UC_{no} and SC pairs, further studies with a whole set of assimilation pairs can fully test the applicability of PAM-S to sequence recall. With an increasing number of non-Chinese students with diverse language backgrounds in Hong Kong, more research is needed to assess their performance in perceptual learning of lexical tone in order to test the theories of language perception and inform pedagogical practice.

6. REFERENCES

[1] Census and Statistics Department. 2016. Thematic report: Ethnic minorities. Hong Kong: Census and Statistics Department. https://www.censtatd.gov.hk/en/data/stat_report/product/B1120100/att/B11201002016XXXXB0100.pdf.

[2] Li, D. C. S., Chuk, J. Y. P. 2015. South Asian students' needs for Cantonese and written Chinese in Hong Kong: A linguistic study. *International Journal of Multilingualism* 12(2), 210–224.

[3] Loh, E. K. Y., Tse, S. K. 2012. An investigation of school-based curriculum of effective Chinese character learning for Non-Chinese speaking kindergarteners. *Han-Character and Classical Writing Language Education* 28, 171–195.

[4] Lee, S.-l. 2019. Teaching and learning Cantonese as a second language: Learners' attitudes and learning hurdles. In: Wakefield, J. C. (ed), *Cantonese as a Second Language: Issues, Experiences and Suggestions for Teaching and Learning*. Routledge, 47–66.

[5] Qin, Z., Mok, P. P. K. 2014. Discrimination of Cantonese tones by speakers of tone and non-tone languages. *Kansas Working Papers in Linguistics* 34, 26–42.

[6] Lee, S.-l. 2020. Learning needs of Cantonese as a second language in Hong Kong tertiary education. *Electronic Journal of Foreign Language Teaching* 17(2), 460–478.

[7] Yu, A. C. L., Lee, C. W. T., Lan, C., Mok, P. P. K. 2022. A new system of Cantonese tones? Tone perception and production in Hong Kong South Asian Cantonese. *Language and Speech* 65(3), 625–649.

[8] Dupoux, E., Peperkamp, S., Sebastián-Gallés, N. 2001. A robust method to study stress "deafness". *The Journal of the Acoustical Society of America* 110(3), 1606–1618.

[9] Dupoux, E., Sebastián-Gallés, N., Navarrete, E., Peperkamp, S. 2008. Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106(2), 682–706.

[10] Choi, W. 2020. The selectivity of musical advantage: Musicians exhibit perceptual advantage for some but not all Cantonese tones. *Music Perception* 37(5), 423–434.

[11] Kim, H., Tremblay, A. 2021. Korean listeners' processing of suprasegmental lexical contrasts in Korean and English: A cue-based transfer approach. *Journal of Phonetics* 87, 101059.

[12] Choi, W., Chiu, M. M. 2022. Why aren't all Cantonese tones equally confusing to English listeners? *Language and Speech*, 00238309221139789.

[13] So, C. K., Best, C. T. 2014. Phonetic influence on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition* 36(2), 195–221.

[14] So, C. K., Best, C. T. 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53(2), 273–293.

[15] Chen, J., Best, C. T., Antoniou, M., Kasisopa, B. 2019. Cognitive factors in perception of Thai tones by naïve Mandarin listeners. In: *Proc. 19th ICPhS Melbourne*, 1684–1688.

[16] Chen, J., Antoniou, M., Best, C. T. 2023. Phonological and phonetic contributions to perception of non-native lexical tones by tone language listeners: Effects of memory load and stimulus variability. *Journal of Phonetics* 96, 101199.

[17] Evans, J., Yeh, W. C., Kulkarni, R. 2018. Acoustics of tone in Indian Punjabi. *Transactions of the Philological Society* 116(3), 509–528.

[18] Nara, K. 2016. Acoustic analysis of Punjabi stress and tone (Doabi dialect). *The Journal of the Acoustical Society of America* 139(4), 2215–2215.

[19] Singh, A., Pandey, D., Agrawal, S. S. 2015. Analysis of Punjabi tonemes. In: *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, 1694–1697.

[20] Bauer, R. S., Benedict, P. K. 1997. *Modern Cantonese Phonology*. Berlin: Walter de Gruyter.

[21] Lix, L. M., Keselman, J. C., Keselman, H. J. 1996. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research* 66(4), 579–619.

[1] ++ = firmly assimilated, + = moderately assimilated.

[2] The error bars show 95% confidence intervals.