THE EFFECT OF NATIVE PHONOLOGY CONSTRAINTS ON THE PERCEPTION OF NON-NATIVE TONES

Jinghong Ning, Yi Liu

The Hong Kong Polytechnic University
chion.ning@polyu.edu.hk, yi.liu@polyu.edu.hk

ABSTRACT

Native context-based phonological constraints have been proven to be influential in acquisition of second language (L2) contrasts on the segmental level. However, it still needs to check whether such first language (L1) phonological constraints also play a role in the perception of L2 suprasegments by non-native speakers. Tonal language speakers (Shanghai Wu dialect), whose L1 obeys a sandhi rule dissimilar to that in Mandarin, served as the experimental group, and Mandarin native speakers served as a control group. In the results, a context effect was detected for the non-native group, which showed the transfer of phonology constraints in Shanghai. Thus, it is suggested that L1 context-based constraints would be incorporated into the framework of PAM when making perceptual predictions across languages.

Keywords: context-based phonological constraints, categorical perception, L2 exposure, PAM

1. INTRODUCTION

1.1. Tonal perception

Generally, linguistic boundaries may exist for native speakers when they perceive phonemes from their own language systems [11]. Cross-linguistically, native speakers show perceptual edges over L2 learners when perceiving their mother tongue, whereas adult non-native speakers have to undergo a long period of language learning to achieve a native-like categorical level [14].

The L2 perception has been proved to be influenced by several factors, among which, L1 experience is documented as one of the main linguistic effects in perceiving L2 phonemes. As one of influential L2 learning models, the Perceptual Assimilation Model (PAM) ([2], [19]) hypothesizes that, non-native phonemes can be assimilated into perceivers' L1 prosody categories. Exemplars exist for which the goodness of fit can range from very good to poor. For an example, if two L2 phonemes are equally mapped to one exemplar in the L1 system, and the resultant perception is predicted to be poor.

Although PAM highlighted the dual effects from both the phonetic and phonological aspects, those effects have been richly tested by a majority of perceptual studies that predominantly emphasized the phonetic level, and weighted heavily toward the phonetic inventory ([9], [11], [15]). Only a few studies took into account the influence of L1 phonology when predicting the perceptual behaviors by non-native listeners ([12], [19]), and barely no research took into consideration of the L1 context-based constraints. For an instance [12], to investigated the effect of L1 sandhi rules in perception of L2 tone contrasts. Subjects from the three dialect-speaking areas and experienced English learners of Mandarin were invited to identify Rugao, Yantai, and Mandarin tones, which were dissimilar in tone sandhi rules. In the results, both sandhi and typology effects were detected, since different perceptual-error types were generated according to different sandhi rules in the speakers' native prosody systems.

On the segmental level, the context-based L1 phonology transfer has been detected both in L2 segment production ([3], [4]) and perception [7]. It is also noteworthy that context-based constraints will be unlikely to spread to all the situations that occur in L2, and instead, it is more likely that only a preserved structure or a similar context can trigger such a transfer [3]. As reported in [4], when acquiring French and Spanish rhotics, English (L1) speakers showed sensitivity to the flap locations, which could hardly be fully explained by the current L2 learning models. Thus, it is necessary to incorporate speech constraints, such as position-based differences and unequally development in phonetic parameters into the current L2 learning models when making predictions of the production difficulty for L2 learners [4].

1.2. L1 phonology and categorical features

1.2.1. Beijing Mandarin

Four primary contrastive lexical tones exist in Mandarin. Under the 5-scale pitch framework [6], the citation forms of the tones are usually given as follows: the high-level tone (T1, 55), the mid-rising tone (T2, 35), the dipping tone/low-falling-rising

tone (T3, 214), and the high-falling tone (T4, 51). Contextually, T3 undergoes a leftward sandhi when followed by another T3 [5].

In terms of Mandarin tones, T2 and T3 are rated as the most confusing pair among the four lexical tones for L1 and L2 listeners [15]. Hao (2012) [9] rated the T2 and T3 confusion to be caused by the natural acoustic similarity between the two tones. Nevertheless, such an L1-independent view has been challenged by some other empirical work ([12], [15]) pointing out that the perceiving of the instinctive confusing pair T2 and T3 could be also impacted by listeners' preexisting phonetic systems.

1.2.2. Shanghai Wu

The contemporary Shanghai Wu dialect has five citation tones [21]: the Yin (upper register) high-falling tone (W1, 53), the Yin mid-rising tone (W2, 34), the Yang (lower register) low-rising tone (W3, 13/14), the Yin high-checked tone (W4, 5), and the Yang low-checked tone (W5, 2/12). There exist context-based sandhi constraints in the prosody system of Shanghai. Specifically, Shanghai undergoes a rightward sandhi when the two syllables are tightly connected in lexical meaning, and a leftward sandhi will occur instead if the two syllables are loosely correlated to each other. Because the right-spreading rule occupies the majority of the daily-use vocabulary, the rightward sandhi seems dominant in the Shanghai dialect [22]. As shown in (1) to (3), there are primarily five constrained allotones in Shanghai, among which the high-level variant (FI, 55) for W1, the mid-level variant (FII, 33) for W2, and the low-level variant (FIII, 22) for W3 emerge only in the first syllables, while all the tones shift to the low-falling variant (SI, 21) and the mid-level variant (SII, 44) in the second syllables.

(1) W1+X→FI+SI; (2) W2+X→FII+SII
(3) W3+X→FIII+SIII

1.3. Current study

PAM theory, on segmental and suprasegmental levels, successfully establishes a bond between the perceptual behaviors and learners' language backgrounds. In applying the PAM, a majority of studies that support PAM tend to explain the phonetic features when establishing a bond between L1 and L2 typologies, perhaps losing sight of the irreplaceable role played by phonological constraints and status. The main objective of this study was to see whether a context-based phonological constraint, which has been reported to be influential on the

segmental level, can impact non-native speakers' tone perception.

2. EXPERIMENT DESIGN

2.1. Participants

Eighteen native speakers of Mandarin from Beijing (nine females and nine males, with a mean age = 20.1 yrs., SD = 2), and 20 Shanghai Wu dialect speakers from a Shanghai urban area (10 females and 10 males) participated in the experiment. All of the subjects were right-handed and had no musical training experience [13]; Also, none of them suffered any speech or hearing difficulty. The dialect speakers used Mandarin in school and spoke dialect in their families. They were born and raised in Shanghai city and had never left their hometown.

2.2. Materials

Real disyllabic words in Mandarin were partially adopted from [16]. The selected words were balanced in reference tone type (high and low) with the target tones located either in the first or second syllables. Each tonal combination contained two words. All disyllabic words were recorded by a male native speaker of Mandarin from Beijing.

Figure 1 Scheme for stimuli synthesis [18]
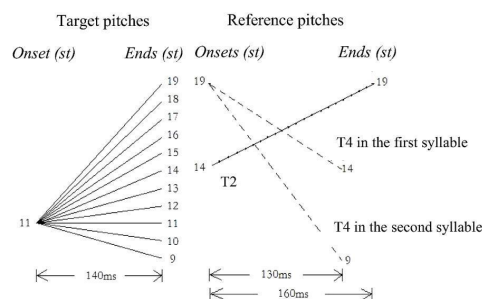
To synthesize the stimuli, pitches were transformed into semitones, and synthesized using the approach of PSOLA [16] in Praat. Given that both native and non-native listeners could take advantage of F0 height to distinguish pitches [8], the pitch onsets of target syllables were set as 11st, and offsets ranged from 9st to 19st with 1 semitone difference. For target syllables, duration was normalized at 140 ms. The reference syllables were normalized as follows: 14-19st, 160 ms for T2, 19-14st; 130 ms for T4 in the first syllable; 19-9st, and 130 ms for T4 in the second syllable. The mean intensity for all syllables was normalized at 65 dB.

2.3. Procedure

The experiment was completed in a sound-treated room at the Hong Kong Polytechnic University for Mandarin native speakers, and at the Shanghai Jiaotong University for the Shanghai group. During the procedure, the participants sat in front of a 13-inch desktop (i5 core, USB interface: 3.0) with a Boom headphone (Nady QHM-100 stereo).

The identification task, were programmed with E-prime 2.1 (Psychology Software Tools, Inc., Pittsburgh, Pa., USA). The binomial forced-choice AX paradigm with an inter-stimulus interval of 1500-ms was employed. The subjects completed the exercise block first, and in an identification task, participants heard one stimulus at a time and recognized it from two given words displayed. The stimuli randomized with 5 repetitions, and the whole experiment was accomplished within one hour.

2.4. Data analysis

Categoricity is commonly described with 50% crossover of identification curves, a linear distance between 25% and 75% percentile (boundary width, CB), and a discrimination peak [20]. The discrimination peak and category position have been demonstrated to be language independent, whereas the CB has been reported to be L1 dependent ([17], [19]). Thus, only CB was further utilized as an indicator for category degree in this report.

The identification rate is defined as the proportion of the occurrence of target tones out of total responses [17]. The CB were generated with Probit analysis [10], and those CB values with significant $\chi2$ ($p < .05$) were filtered out from the data pool. Eventually, 140 (18 subjects × 8 words - 4 outliers) CB values were acquired for the Mandarin group and 150 (20 subjects × 8 words - 10 outliers) CB values were collected for the Shanghai group.

LMM were applied with lme4 package in R (ver. 3.5.1) due to their clear advantages over multiple regression in assessing hierarchical, nested data [1]. Likelihood Ratio Tests (LRT) were conducted to decide the intercepts and slopes of the random effects. Bonferroni-correction was adopted to adjust p-values for multiple comparisons.

3. RESULTS

An LMM was conducted with CB as the dependent variable, and with L1 background (Mandarin, Shanghai), syllable location (first and second syllable), and reference tone (High, Low) as fixed factors. LRT ($\chi2$ (2) = 1.22, p = .54) suggested that subject slopes for location and reference tone were not included, and only subject and word intercepts

were considered as random effects. Since no significance was reported by tone reference, it was thus removed from the model. According to the simplified LMM, for the random effects, the subject intercept exhibited a random effect with a variance of 0.003 (SD = 0.06), and the word intercept had a variance of 0.012 (SD = 0.11).

Table 1 Parameters of LMM for CB results. Estimate (β), standard errors (SE), degree of freedom (df), t value and p value were reported for the fixed effects.

	β	SE	df	t	p
Intercept	1.26	0.28	69.9	4.43	<.001
L1	0.09	0.16	280.4	0.6	.54
Location	-0.3	0.18	68.6	-1.8	.071
L1× Location	0.31	0.1	246.2	3.08	.002

The average values of CB were exhibited in Figure 2 in terms of the different syllable locations and reference tones. According to figure 2, firstly, the CB values for Mandarin participants narrowly ranged from 1.24 to 1.44, whereas, the values fell in a large range of 1.73 to 2.21 in terms of the Shanghai subjects. Thus, a clear L1 difference of CB can be seen between the native and non-native groups.

Moreover, with regard to the fixed effects,which were illustrated in Table 1, an interaction (β = 0.31, SE = 0.1, t (246.2) = 3.08, p = .002) between the L1 background and syllable location was found in terms of CB, indicating a location difference for certain L1 groups. Hence, the contextual differences in CB would be further uncovered depending on different L1 groups.

Figure 2 Bar graph for boundary width across Mandarin and Shanghai speakers in terms of different syllable locations and tone references. The standard deviations are illustrated with vertical bars.
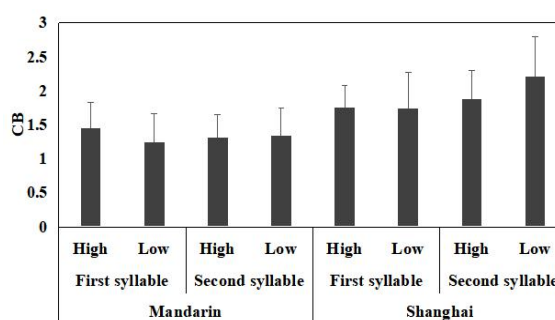
Mandarin group, according to post-hoc Tukey test, was free of any syllable location effect (p = .93), with a CB value of 1.33 (0.41) in the first syllable

and 1.32 (0.37) in the second syllable. However, a location difference was also detected in Shanghai group, with Shanghai CB value was averagely much higher ($p < .001$) in the second syllable (mean = 2.06, SD = 0.54) than in the first syllable (mean = 1.75, SD = 0.42). In other words, compared with the native speakers, the Shanghai subjects showed a clear location-specific perceptual performance.

4. DISCUSSION AND CONCLUSION

The smaller the CB value, the higher the degree of categorization that will be obtained toward the sound continuum [11]. In the current results, Mandarin speakers showed a native perceptual advantage over Shanghai group, with narrow category boundaries for the identification task. Thus, the Mandarin result confirms the declaration of an "L1 advantage" [14].

Moreover, it was interesting to find that the contextual environment played different roles in the perceptual results of native and non-native speakers. Mandarin speakers behaved insensitively toward the different target locations, which is consistent with the previous findings [18], who used stimuli that were similar to the ones we used in the experiment. In contrast, the Shanghai speakers had a noticeably higher categoricity when the tones were targeted at the first syllable location. Hence, we were eager to see why the contextual effect in the Shanghai group drifted away from that in the native group. A contextual difference in non-native sound perception is probably derived from learning strategy or phonology impact [4]. According to the previous study [9], experienced English-speaking learners of Mandarin were likely to have perceptual advantages when the target sound was located in the second syllable position, and learners showed their highest error rate for the first position. Hence, the contextual effect for Shanghai speakers does not tend to result from the application of L2 learning strategies. On the contrary, it might attribute to the phonological transfer of speakers' L1 prosody system instead. Thus, the current work expands the previous studies ([3], [7]) to a suprasegmental dimension, emphasizing that the L1 context-based constraints might impact the dialect speakers' perception of non-native tone contrasts. Moreover, the current study also supports the claim [12] that for dialect speakers, L1 and L2 tone systems are not stored in separation but instead, they also fit non-native tones into the existent native norms, such as the tone sandhi rule.

Accounting for PAM, a large amount of previous research attached great importance to the effect of L1 typology. Both Mandarin T2 and T3 are assimilated as W3 tone category in Shanghai [21], so a poor perceptual result is expected for Shanghai speakers when perceiving the two tones, which goes along with the current results. However, the contextual effect detected for Shanghai speakers cannot be fully explained by such typology-based prediction. Since Shanghai dialect undergoes a context-based tone sandhi, the low-rising W3 actually realizes as either high, mid or low level variants in the first syllables, and shifts to mid-level or low-falling variants when located in the second syllables. Instead of a single assimilation pattern introduced in [21], different tone assimilation might occur according to syllable locations, leading to different perceptual predictions. This is because the transfer of context-based constraints only occurs in a similar structure instead of spreading to all environments ([3], [7]). In the near future, a follow-up experiment is expected to be done to explore the assimilation pattern between Mandarin and Shanghai tone categories in terms of different syllable locations.

Similarly, when investigating the production of the Spanish alveolar taps by English speakers, Colantoni and Steele (2008) [4] found that the results varied vastly depended on the flap locations being under the impact of the position-based constraints of rhotics in English prosody. Together with [4] , the current study thereby supports the proposal that the explanatory power of PAM will be enlarged if the effect of the L1 context-based constraints are incorporated into the model's theoretical framework.

This article identifies the impact that L1 phonological constraints exerted on the acquisition of L2 tone contrasts. A contextual difference, which was detected in the perception results for the Shanghai speakers, cannot be successfully predicted by a typology-based assimilation pattern with respect to PAM. And it is necessary to incorporate constraint effects into the framework of the PAM theory on a suprasegmental level. Beyond PAM, those results reinforce and extend the existing view about the effect of L1 phonological constraints, from a segmental level to a suprasegmental level.

5. ACKNOWLEDGMENT

6. REFERENCES

[1] Baayen, R. H., Davidson, D. J., Bates, D. M. 2008. Mixed-effects modeling with crossed random effects

for subjects and items. *Journal of Memory and Language*, 59, 390–412.

[2] Best, C. T. 1995. A direct realist view of cross-language speech perception. In W. Strange (eds.), *Speech Perception and Linguistic Experience Issues in Cross- Language Research*. Baltimore: York Press, 171-204.

[3] Carlisle, R. 1998. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition*, 20, 245–260.

[4] Colantoni, L., Steele, J. 2008. Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics*, 29, 489-534.

[5] Chao, Y. R. 1968. Interlingual and interdialectal borrowings in Chinese. *Shock*, 15, 19-28.

[6] Chao, Y. R. 1980. A system of tone letters. *La Maître phonétique*. 45, 24- 27.

[7] Eckman, F. R., Iverson, G. K., Fox, R., Jacewicz, E., Lee, S. 2009. Perception and production in the acquisition of L2 phonemic contrasts. In M. A. Watkins, A. S. Rauber, and B. O. Baptista (eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production.* Cambridge: Cambridge Scholars Publishing, 81-95.

[8] Gandour, J. 1978. The perception of tone. In V. A. Fromkin (eds.), *Tone: A Linguistics Survey*. New York: Academic Press, 41-76.

[9] Hao, Y. C. 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40, 269-279.

[10] Hallé, P. A., Best, C. T., Levitt, A. 1999. Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, 27, 281-306.

[11] Hallé, P. A., Chang, Y. C., Best, C. T. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32, 395-421.

[12] Huang, T. 2012. Cross-linguistic and inter-dialectal differences in tone perception by native speakers of three Chinese dialects and American English. *Journal of Chinese linguistics*, 40, 155-179.

[13] Kiriloff, C. 1969. On the auditory perception of tones in Mandarin. *Phonetica*, 20, 63-67.

[14] Lee, Y. S., Vakoch, D. A., Wurm, L. H. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25, 527-542.

[15] Leung, A. 2008. Tonal assimilation patterns of Cantonese L2 speakers of Mandarin in the perception and production of Mandarin tones. *Proc. of the 2008 CLA Annual Conference.*

[16] Moulines, E., Laroche, J. 1995. Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Commun.*, 16, 175-205.

[17] Peng, G., Deutsch, D., Henthorn, T., Su, D., Wang, S. Y. 2013. Language experience influences nonlinguistic pitch perception. *Journal of Chinese Linguistics*, 41, 447-467.

[18] Rong, R., Wang, P., Liang, L., Shi, F. 2015. Hanyu putonghua shengdiao de tingjue ganzhi geju [The perception pattern in Mandarin]. *Nankai Yuyan Xuekan [Nankai Linguistics]*, 1, 11-24.

[19] 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, 273-93.

[20] Treisman, M., Faulkner, A., Naish, P. L., Rosner, B. S. 1995. Voice-onset time and tone-onset time: The role of criterion-setting mechanisms in categorical perception. *Quarterly Journal of Experimental Psychology*, A, 48, 334-366.

[21] Xu, B. H., Tang, Z. Z., You, R. J., Qian, N. R., Shi, R. J., Shen, Y. M. 1988. *Shanghai fangyanzhi [Records of Shanghai dialect]*. Shanghai: Shanghai Jiaoyu Chubanshe [Shanghai Educational Publishing House], 124-125.

[22] Zhu, X. N. 2005. *Shanghai shengdiao shiyanlu [Shanghai tone experimental record]*. Shanghai: Shanghai Jiaoyu Chubanshe [Shanghai Educational Publishing House], 233-239.