INFLUENCES OF MULTI-LANGUAGE EXPERIENCE ON MANDARIN CONSONANT PERCEPTION

Shuwen Chen[1], Janice Wing Sze Wong[2]

[1]Institute of Linguistics, Chinese Academy of Social Sciences, [2]Hong Kong Baptist University
[1]chensw@cass.org.cn, [2]janicewong@hkbu.edu.hk

ABSTRACT

Cantonese speakers from Hong Kong learn Mandarin under the influence of both Cantonese and English. This study investigates the influence of Cantonese and English on Mandarin consonant perception. Nineteen Cantonese speakers were tested in an identification task followed by a cross-language mapping and goodness-rating task when perceiving Mandarin fricatives, affricates, liquids and the alveolar nasal. The results showed that Cantonese-English bilinguals mapped Mandarin consonants onto both Cantonese and English consonants. Language dominance and perceptual similarity both play a role in cross-language assimilation. The cross-language mapping pattern, however, only partially predicts the identification results. In addition, mapping Mandarin consonants onto English sound categories has a facilitative effect on the perception of Mandarin consonants.

Keywords: Speech perception, Mandarin consonants, Cantonese-English bilinguals

1. INTRODUCTION

An important question in L2 speech acquisition is how cross-language similarities and differences affect and predict the acquisition of L2 sounds. Perceptual Assimilation Model - L2 (PAM-L2) [1] predicts that if two non-native sounds are assimilated to two different L1 categories, the perception of the two sounds will be good. If two non-native sounds are assimilated to the same L1 category, the discrimination of the two sounds will be poor. If two non-native sounds are assimilated as a good version and a poor version of the same category, the accuracy is predicted to be intermediate. Previous studies on L2 vowel and consonant perception have provided some support for this hypothesis [2, 3]. A relevant and intriguing question is whether the link between cross-language similarities and perception results can be extended to the scenario where a third sound system is learned. Previous studies suggested that the first two sound systems can both have an impact on the production and perception of the third sound system [4, 5]. But will the sound categories in the third system be mapped onto the sound categories in the first or the second sound system? Can cross-language mapping patterns predict perceptual accuracy when three sound systems interact?

To answer these questions, the current study examined the perception of Mandarin consonants by Cantonese-English bilinguals. Cantonese and Mandarin are both Chinese languages, but their phonological systems are quite different. Mandarin has a rich inventory of sibilants. It contrasts three manners of articulation for dental sounds (/s, ts, ts^h/), retroflex sounds (/ʂ, tʂ, tʂ^h/) and alveolo-palatal sounds (/ɕ, tɕ, tɕ^h/). In contrast, Cantonese has a much smaller inventory of sibilants (/s, ts, ts^h/). English has six alveolar and post-alveolar sibilants (/s, z, ʃ, ʒ, tʃ, dʒ/). In addition, both Mandarin and English contrast /n/, /ɹ/, and /l/, while Cantonese does not have a /ɹ/ phoneme and is gradually losing /n/-/l/ contrast [6]. Given the differences in consonant inventories, it is interesting to examine how Cantonese-English bilinguals map Mandarin consonants onto Cantonese or English consonants. Does English facilitate Mandarin consonant perception? Can cross-language assimilation patterns predict the accuracy of Mandarin consonant identification?

2. METHOD

2.1. Participants

Nineteen undergraduate students (15 females and 4 males) from Hong Kong Baptist University were invited to participate in the experiment. Their average age was 20 years old. They were born and grew up in Hong Kong. They all speak Cantonese as their first language, and learned English and Mandarin later. They started learning English at an average age of three and Mandarin at an average age of four. In the language background survey, participants reported that they used Cantonese and English more frequently than Mandarin in daily communication and in school, and had a higher proficiency in Cantonese and English than in Mandarin. All participants formally learned *Pinyin* before participating in the experiment.

2.2. Stimuli

The target consonants included Mandarin alveolar fricative and affricates /s, ts, ts^h/, retroflex fricative

and affricates /ʂ, tʂ, tʂʰ/, and palatal fricative and affricates /ɕ, tɕ, tɕʰ/, approximants /l, ɹ/ and alveolar nasal /n/. The fricatives and affricates are followed by the high front vowel /i/ or the apical vowel /ɿ, ʅ/, and the liquids and the nasal are followed by /i/ and /a/ (see Table 1). The stimuli were read by two Mandarin native speakers from Beijing (one male and one female) and recorded in a sound-treated room. The target syllables were embedded in the carrier phrase "wo ʂwo __ tʂɤ kɤ tsʅ" (I say ___ this word). After the recording, the target syllables were cut out to be used in the experiment.

Target consonant	Stimuli	Target consonant	Stimuli
ts	tsʅ	n	ni
tsʰ	tsʰʅ		nan
s	sʅ	l	li
tʂ	tʂʅ		lan
tʂʰ	tʂʰʅ	ɹ	ɹʅ
ʂ	ʂʅ		ɹan
tɕ	tɕi		
tɕʰ	tɕʰi		
ɕ	ɕi		

Table 1: Stimuli used in the experiment.

2.3. Procedure

The participants completed an identification task, and a cross-language mapping and goodness rating task in a classroom. The stimuli were presented through the loudspeakers in the classroom, and the subjects were asked to respond using online questionnaires. The stimuli were played in a random order using Praat [7]. In the identification task, the subjects heard the stimuli and they were asked to choose which sound it was in *Pinyin*. The 30 stimuli (15 words × 2 speakers) were repeated three times, resulting in a total of 90 trials. In the second task, participants listened to the stimuli and were asked to classify the sounds they heard as one of the Cantonese categories (/s, ts, tsʰ/) or English categories (/s, z, ʃ, tʃ, dʒ/). English /ʒ/ was not included because it cannot occur in the syllable initial position. They were then asked to rate the similarity between the word they heard and the sound they chose on a scale of 1 (unlike) to 5 (very similar). They were allowed to choose one sound (either Cantonese or English), or two sounds (one from Cantonese and one from English). Each stimulus was identified and rated two times, resulting in a total of 60 trials.

2.4. Statistical analysis

For the identification results, identification accuracy for each sound was calculated, and confusion matrices were summarized. For the cross-language mapping and goodness rating task, the percentage of identification for each category in Cantonese and English was calculated. Following Guion et al. [8], the "fit index" was calculated to combine the identification percentage and goodness rating into one single matrix. It is calculated by multiplying the percentage of identification by the goodness rating of that identification. For example, if the percentage of identifying Mandarin /ts/ as Cantonese /ts/ is 24%, and the goodness rating of this identification is 1.67, then the fix index of this classification is 0.40.

3. RESULTS

3.1. Identification results

Figure 1 shows the accuracy of identifying Mandarin consonants by Cantonese-English bilinguals. Mandarin /ɹ/ and /n/ were identified with the highest accuracy. The accuracy of /tɕʰ/ (88%) and /tɕ/ (87%) was highest among the fricatives and affricates. The accuracy of /tsʰ/ (62%) and /s/ (71%) was the lowest among all the tested consonants.

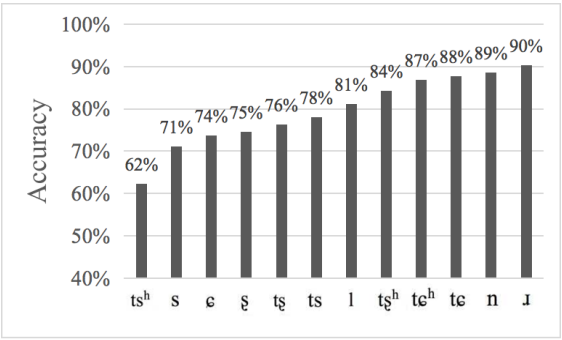

Figure 1: Accuracy of identifying Mandarin consonants

Table 2 is the confusion matrix of Mandarin fricatives and affricates. First, it is easy for the speakers to confuse Mandarin /ɕ/ and /s/. It is also hard for them to distinguish the three retroflex sounds / ʂ, tʂ, tʂʰ/, where /tʂ/ and /ʂ/ were often identified as /tʂʰ/. In addition, Mandarin /tsʰ/ was often identified as /tɕʰ/, /tʂʰ/ or /ts/.

		Responses								
		ts	tsʰ	s	tʂ	tʂʰ	ʂ	tɕ	tɕʰ	ɕ
Stimuli	ts	89	5	1	7	0	0	8	3	1
	tsʰ	10	71	1	0	10	0	2	18	2
	s	2	10	81	2	0	3	0	2	14
	tʂ	5	1	1	87	14	0	5	1	0
	tʂʰ	0	5	0	5	96	4	0	4	0
	ʂ	0	2	6	3	16	85	0	0	2
	tɕ	8	3	0	1	1	0	100	1	0
	tɕʰ	1	5	2	1	6	0	0	99	0
	ɕ	1	7	15	1	0	3	0	3	84

Table 2: Confusion matrix of Mandarin fricatives and affricates.

Table 3 is the confusion matrix of Mandarin nasal /n/ and approximants /l, ɹ/. The most common confusion was between /n/ and /l/. The confusion between /l/ and /ɹ/ can also be seen, but was less frequent.

		Responses		
		ɹ	l	n
Stimuli	ɹ	206	13	9
	l	12	185	31
	n	3	23	202

Table 3: Confusion matrix of Mandarin nasal and approximants.

3.2. Perceptual assimilation results

Table 4 and 5 summarized the mean percent identification and goodness rating of Mandarin consonants in terms of Cantonese and English sound categories. When compared to Cantonese categories, Mandarin /ts/, /tʂ/ and /tɕ/ were most similar to Cantonese /ts/, Mandarin /tsʰ/, /tʂʰ/ and /tɕʰ/ were most similar to Cantonese /tsʰ/, and Mandarin /s/ were most similar to Cantonese fricative /s/. When compared to English categories, Mandarin /s/ and /ɕ/ were assimilated to English /s/, Mandarin /ts/ to English /z/, Mandarin /ʂ/ to English /ʃ/, Mandarin /tsʰ/, /tʂʰ/ and /tɕʰ/ to English /tʃ/, Mandarin /tʂ/ and /tɕ/ to English /dʒ/.

As to the liquids and the nasal, Mandarin /l, ɹ, n/ were mapped onto English /l, ɹ, n/ respectively. When compared to Cantonese sounds, Mandarin /l, ɹ, n/ were all assimilated to Cantonese /n/.

	Cantonese			English			
	n	l	None of them	l	n	ɹ	None of them
l	**8 (1.38)**	4 (2)	64	**26 (3.04)**	10 (2.6)	1 (2)	39
n	**9 (1.44)**	3 (2.67)	64	3 (2.33)	**33 (3.06)**	1 (2)	39
ɹ	**3 (1.67)**	2 (3)	71	1 (1)	3 (3)	**9 (1.33)**	63

Table 5: Mean percent identification and goodness rating (in parentheses) of Mandarin nasal and approximants in terms of Cantonese and English categories. Boldfaced values indicate the most commonly identified responses in each language.

To compare the cross-language mapping pattern more directly, the fit index of the most commonly identified Cantonese and English sounds was calculated and summarized in Table 6. A higher score of fix index indicates more similarity between Mandarin consonants and the identified sound. According to Table 6, among the 12 Mandarin consonants, 6 were mapped onto Cantonese consonants, and 6 were mapped onto English consonants. Mandarin /tɕ/ and /tʂ/ were mapped onto Cantonese /ts/, Mandarin /s/ and /ɕ/ were mapped onto Cantonese /s/, Mandarin /tsʰ/ and /tɕʰ/ were mapped onto Cantonese /tsʰ/. The rest of the Mandarin consonants were mapped onto English sounds. Mandarin /tʂʰ/ and /ʂ/ were mapped onto English /tʃ/ and /ʃ/ respectively. As expected, Mandarin /n, ɹ, l / were mapped onto English /n, ɹ, l/. But unexpectedly, Mandarin /ts/ was mapped onto English /z/ rather than Cantonese /ts/.

	Cantonese				English					
	ts	tsʰ	s	None of them	s	z	ʃ	tʃ	dʒ	None of them
ts	**18 (1.67)**	0	2 (2)	56	0	**24 (2)**	0	2 (2)	6 (1.67)	44
tsʰ	3 (1.33)	**19 (1.42)**	0	54	1 (1)	4 (2.5)	0	**9 (1.78)**	1 (1)	61
s	2 (1.5)	2 (2.5)	**21 (2.14)**	51	**9 (2)**	6 (2.67)	1 (2)	1 (1)	0	59
tʂ	**21 (2)**	0	1 (1)	54	0	5 (1.4)	0	4 (2)	**9 (1.78)**	58
tʂʰ	1 (1)	**17 (1.82)**	1 (2)	57	1 (3)	4 (2)	0	**16 (2.06)**	0	55
ʂ	0	6 (2.17)	**7 (1.14)**	63	1 (1)	1 (1)	**14 (1.43)**	9 (1.78)	1 (1)	50
tɕ	**56 (3.48)**	5 (2.4)	3 (1.33)	12	3 (2)	21 (3.43)	0	4 (2.75)	**31 (2.97)**	17
tɕʰ	3 (2.67)	**54 (3.35)**	1 (1)	18	1 (2)	8 (3)	1 (1)	**23 (2.65)**	2 (2.5)	41
ɕ	1 (4)	2 (1.5)	**51 (2.96)**	22	**47 (3.15)**	6 (2.5)	1 (3)	1 (1)	0	21

Table 4: Mean percent identification and goodness rating (in parentheses) of Mandarin fricatives and affricates in terms of Cantonese and English categories. Boldfaced values indicate the most commonly identified responses in each language.

		Cantonese				English				
		Most common identification	Proportion of identification	Goodness rating	Fit index	Most common identification	Proportion of identification	Goodness rating	Fit index	Most fit sound category
Mandarin consonants	ts	ts	24%	1.67	0.40	z	32%	2.00	0.64	English /z/
	tsʰ	tsʰ	25%	1.42	0.36	tʃ	12%	1.78	0.21	Cantonese /tsʰ/
	s	s	28%	2.14	0.60	s	12%	2.00	0.24	Cantonese /s/
	tʂ	ts	28%	2.00	0.56	dʒ	12%	1.78	0.21	Cantonese /ts/
	tʂʰ	tsʰ	22%	1.82	0.40	tʃ	21%	2.06	0.43	English /tʃ/
	ʂ	s	9%	1.14	0.10	ʃ	18%	1.43	0.26	English /ʃ/
	tɕ	ts	74%	3.48	2.58	dʒ	41%	2.97	1.22	Cantonese /ts/
	tɕʰ	tsʰ	71%	3.35	2.38	tʃ	30%	2.65	0.80	Cantonese /tsʰ/
	ɕ	s	67%	2.96	1.98	s	62%	3.15	1.95	Cantonese /s/
	l	n	11%	1.38	0.15	l	34%	3.04	1.03	English /l/
	n	n	12%	1.44	0.17	n	43%	3.06	1.32	English /n/
	ɹ	n	4%	1.67	0.07	ɹ	12%	1.33	0.16	English /ɹ/

Table 6: Fit indexes for Mandarin consonants in terms of Cantonese and English categories.

4. DISCUSSION

Based on the results of the cross-language mapping task, the listeners map Mandarin consonants onto both Cantonese and English sound categories. This is consistent with Wrembel et al. [9] which shows that the phonemes in the third sound system are mapped onto categories in the other two sound systems. Our results also suggest that both language dominance and perceptual similarity play a role in cross-language mapping. Although Cantonese has a much smaller inventory of sibilants than English does, the bilinguals still map six Mandarin sibilants onto Cantonese, and mapped the other three Mandarin sibilants onto three English sibilants. It should also be noted that in the perceptual assimilation task, the majority of responses were "none of them" for many consonants. It seems to indicate that the participants are aware of the cross-linguistic difference between the three sound systems, and try to keep them separate.

The cross-language mapping pattern only partially predicts Mandarin identification results. According to PAM-L2, if two sounds are mapped onto the same L1 category, it would be hard for learners to distinguish the two sounds. According to the mapping results, Mandarin /s/ and /ɕ/ were both mapped onto Cantonese /s/, and Mandarin /tsʰ/ and /tɕʰ/ were both mapped onto Cantonese /tsʰ/. In the identification task, we did find that participants often mixed /s/ with /ɕ/, and /tsʰ/ with /tɕʰ/. However, there are some exceptions. Mandarin /tɕ/ and /tʂ/ were both mapped onto Cantonese /ts/, but Mandarin /tɕ/ were seldom mixed with /tʂ/, and were often identified as /ts/. Mandarin /tʂ/ was more often identified as /tʂʰ/ than /tɕ/.

In terms of identification accuracy, it is expected that Mandarin consonants that are mapped onto the same Cantonese or English category would have lower accuracy because learners tend to mix those consonants, while consonants that are mapped onto separate categories would have higher accuracy. As expected, we did find that the identification accuracy of Mandarin /ɹ/ and /n/ was high, and Mandarin /s/ and /ɕ/ had a low accuracy. Contrary to the prediction, the identification accuracy of Mandarin /tɕ/ and /tɕʰ/ was high although they were mapped onto Cantonese /ts/ and /tsʰ/ with other Mandarin sounds. It suggests that the relationship between cross-language mapping and perception accuracy is more complex than PAM predicts when there are three sound systems.

The results also show that English consonants facilitate the discrimination of Mandarin consonants for Cantonese-English bilinguals. Cantonese does not contrast /l/ and /ɹ/, but the learners showed high accuracy in identifying the two phonemes. The cross-language mapping pattern has shown that this is because the bilinguals mapped Mandarin /l/ and /ɹ/ onto English /l/ and /ɹ/. A similar effect can be found in /n/-/l/ contrast. Many speakers do not distinguish Cantonese /l/ and /n/ in daily communication, but the identification accuracy of /n/ is still high. This might be the result of a positive transfer from English. This facilitative effect is also reported in Onishi [10].

5. CONCLUSION

In the current study, the perception of Mandarin consonants by Cantonese-English bilinguals was tested. The results showed that the bilinguals mapped Mandarin consonants onto both Cantonese and English categories. Cross-language assimilation is influenced probably by both perceptual similarity and language dominance. Mapping Mandarin sounds onto English sound categories can facilitate the identification of Mandarin consonants when English has contrasts that Cantonese does not have. However, cross-language mapping patterns only partially predict the identification results. It suggests that the relationship between cross-language assimilation and perceptual accuracy is more complex when there are three sound systems.

7. REFERENCES

[1] Best, C. T., Tyler M. D. 2007. Non-native and second language speech perception: Commonalities and complementaries. In Munro, M., Bohn, O-S. (eds), *Second Language Speech Learning*. Amsterdam: John Benjamins Publishing, 13-34.

[2] Tyler, M., Best, C., Faber., A, Levitt, A. 2013. Perceptual assimilation and discrimination of non-native vowel contrast. *Phonetica* 71, 4-21.

[3] Best, C. T., McRoberts, G. W., & Goodell, E. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *J. Acoust. Soc. Am*. 109(2), 775–794.

[4] Zhu, Y., Chen, A., Sudhoff, S., Mok, P.P. 2019. Third language prosody: Evidence from Cantonese-English-German Trilinguals. *Proc. 19th ICPhS*, Melbourne.

[5] Balas, A., Kopečková, R., Wrembel, M., Mickiewicz, A. 2019. Perception of rhotics by multilingual children. *Proc. 19th ICPhS*, Melbourne.

[6] Yip, V., Matthews, S. 2000. *Basic Cantonese: A Grammar and Workbook*. London: Routledge.

[7] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer. http://www.praat.org/.

[8] Guion, S.G., Flege, J.E., Akahane-Yamada, R., Pruitt, J. 2000. An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *J. Acoust. Soc. Am*. 107(5), 2711-2724.

[9] Wrembel, M., Marecka, M., Kopečková, R. 2019. Extending perceptual assimilation model to L3 phonological acquisition. *International Journal of Multilingualism* 16(4), 513-533.

[10] Onishi, H. 2016. The effects of L2 experience on L3 perception. *International Journal of Multilingualism* 13(4), 459-475.