A preliminary study on Cantonese vowel development by young heritage children

Chen Lan, Peggy Mok

The Chinese University of Hong Kong
lchsapphire@gmail.com, peggymok@cuhk.edu.hk

ABSTRACT

This study investigated the developmental pattern of Cantonese vowel system by young heritage speakers (HSs). Eighteen HSs from Vancouver aged 2;0, 4;0, and 6;0 participated in a picture-naming task; seven long vowels and four short vowels were extracted and measured. Results of acoustic analysis suggested that there was no clear distinction among the vowels produced by HSs at age 2;0, displaying a pattern of neutralization to the central position of the vowel space. Vowels became more distributed as age increased, but 6-year-old HSs' production was still more centralized compared to the reference pattern. Further comparison regarding the acquisition of vowel length contrast among HSs, homeland children, and native teenagers indicated that while reference speakers distinguish [aː] vs. [ɐ] in both vowel quantity and quality in all the environments (diphthong and close syllable), homeland children and HSs made use of the cues inconsistently according to environments.

Keywords: heritage language, heritage speakers, Cantonese vowel, language acquisition

1. INTRODUCTION

Heritage language (HL) is a newly recognized form of bilingualism, which has attracted growing research interest in recent years. Heritage speakers (HSs) hear and speak their HL as well as the majority language sequentially or simultaneously in early childhood. Several studies investigated HL phonology in Chinese regarding segments and tones, but most of them were on Mandarin (e.g. [1, 2]) and primarily focused on teenagers or adult speakers (e.g. [3, 4]). One published research investigated young children at the early stages of language development, and it was on Cantonese tones [5]. Little is known about the acquisition of segments in Cantonese by young HSs.

The Government of Canada reports that 208,935 people emigrated from Hong Kong to Canada during the period of 1980 – 2016, and Cantonese is the second most spoken first language among the non-official languages in Canada [6, 7]. The most recent census indicates that 191,940 of 565,270 Canadians who speak Cantonese as their mother tongue are in Vancouver, providing a valuable pool of potential participants for studies on the acquisition of Cantonese by HSs [7]. This study made use of the Vancouver data pool to investigate the production of Cantonese vowels by young HSs, aiming to explore the developmental pattern of vowels and compared it to those reported among the monolingual speakers in Hong Kong.

1.1. Acquisition of Cantonese vowels

The Cantonese vowel system contains seven long vowels [iː yː uː ɛː œː ɔː aː] and six short vowels [ɪ ʊ e ø o ɐ] [8, 9]. Considering [e] and [ɛː] as well as [o] and [ɔː] are variants of the same phonemes, this study analysed them as /ɛː/ and /ɔː/ respectively. Vowels in Cantonese contrast in length and lip rounding. Vowel quality (formant) and quantity (duration) are the significant cues for the distinction of vowel length contrasts. Considering that only the /aː/ vs /ɐ/ pair contrasts in all environments (in diphthongs, before stops and nasal codas) with a high functional load [10, 11], this study only examined this contrastive vowel pair. Regarding vowel roundedness, Cantonese has two pairs of vowels differing in lip roundedness: /iː/ vs. /yː/ and /ɛː/ vs. /œː/.

Previous studies on the acquisition of Cantonese vowels suggested that monolingual children are able to correctly produce most of the vowels and use them contrastively by age 2;3, and the acquisition is completed at age 5;0 [12, 13, 14]. Bilingual pattern was comparable to the monolingual norm [15]. Adult HSs aged between 20 and 87 maintained the 7 vowel contrasts while having inter-generational vowel shift (i.e., the retraction of [yː] and the fronting of [iː]) [16]. Vowel changes were also observed in adult Cantonese speakers (i.e., lowering of [ɪ] and [ʊ]) [16].

1.2. Phonological development of heritage language

To date, there is no framework that specifically targets the speech acquisition of young HSs. A developing predictive model of HL competence on sound system suggested that the phonologically distinct contrasts in HL should be maintained or even overregularized by HSs, and HSs may not simply develop a merged system [17, 18]. Based on this preliminary model, it can be predicted that the HL sound system will be less ambiguous, more regular, and with less innovative merging phenomenon. The Speech Learning Model (SLM [19]) was also utilized to predict the development of a HL sound system, proposing that the interaction between the sound system of the majority language and the HL may contribute to an HL sound system differing from the monolingual norm [20]. Previous studies on HSs supported that the productions of HL segments by HSs, even at pre-schooler age, were significantly different from those of monolinguals [e.g., 21]. In addition, HL acquisition may relate to language variation and change [16]. For instance, the retraction of Cantonese [yː] found in adult HSs did not exist among Hong Kong Cantonese speakers.

Based on these models, we predicted that the development of Cantonese vowels by young HSs may differed from that of their monolingual peers, with a slower acquisition progress. Also, some distinctive vowel contrasts in Cantonese may be overregularized by HSs.

2. METHOD

2.1. Participants

Eighteen HSs aged 2;0 (6, 3f), 4;0 (6, 2f), and 6;0 (6, 4f) who were born and raised in Vancouver, Canada, participated in the production experiment. The age of each participant is more or less no more than 5 months than age 2;0, 4;0, and 6;0. All the participants reported having no hearing or speech problems. Formant frequency data of native female speakers aged between 18 and 21 reported in [9] was used as the reference for comparison. For the acquisition of vowel length contrast, eighteen Hong Kong Cantonese children (homeland children) aged 2;0 (6, 2f), 4;0 (6, 4f), and 6;0 (6, 3f) and nine native speakers (7f) aged between 15 and 18 from [22] were used for comparison.

2.2. Materials

The vowel production experiment was a picture-naming task, which was adopted from the Hong Kong Cantonese Articulation Test (HKCAT) [23], a standardized Cantonese articulation test with norm reference on the acquisition of consonants, vowels, and tones in Hong Kong. All words, such as “faa1 *flower*, ce1 *car*, so1 *comb*, fu3 *a pair of pants*, jyu5 *rain*, daai6 *big*”, were familiar for Cantonese-speaking children. Fifty-two colored pictures, each accompanied by its name in Chinese character(s), were used to elicit 63 Cantonese words. Fourteen of the 63 words were chosen to investigate the 7 long vowels in open syllable and closed syllables, 4 were used to examine the 4 short vowels in closed syllables, and 2 with diphthong were selected to explore the participants’ ability to produce the vowel length contrast (/aː/ vs. /ɐ/).

2.3. Procedure

All participants were paid to do the experiment in a quiet room with parental consents. A detailed language background questionnaire was filled in by the parents of the children. An experimenter, who was a native speaker of Cantonese with phonetical training, facilitated the test. Following the instructions in HKCAT, the experimenter elicited the production by asking questions such as “*ni1 go3 mat1 ye5 lei4 gaa3? What is this?*” while pointing to the picture. Children were instructed to name the target words in isolation twice. No time restriction was imposed. If the participants really could not produce the item, they would repeat the target word after the experimenter. The two clearest repetitions of each item were chosen for analysis. In total, each participants produced 40 words for comparison.

2.4. Data analysis

The acoustic measurements were done in Praat using FormantPro [24]. The first, second and the third formant frequencies of all the tokens were measured, and the average values of 11 vowels were plotted on the F1-F2 and F1-F3 panel. Further evaluation on the vowel length contrast (/aː/ vs. /ɐ/) involved both vowel quantity and vowel quality. The duration of each target vowel production was extracted, measured in millisecond (ms), and reported in second (s). The formant frequencies of the relevant tokens were measured at 10 equidistant points along the vowel interval. Manual check and correction were done. The raw data of formant frequencies in Hz were converted to Bark for normalization [25].

3. RESULTS

3.1. Vowel space

Figure 1 displays the positions of the 7 long vowels produced by the HSs at ages 2;0, 4;0, and 6;0 relative to the reference positions in the F1-F2 panel. In general, there is no clear distinction among the vowels produced by HSs at age 2;0, displaying a pattern of neutralization to the central position of the vowel space. Vowels became more distributed as age increased, but 6-year-old HSs’ production was still more centralized compared to the reference pattern. The high vowels [iː yː uː] and the low vowel [aː] were acquired first at age 2;0, with significant difference in F1 respectively between [aː] and [iː] ($p < .001$), [yː] ($p < .001$), and [uː] ($p < .001$). There was no significant F1 difference between the low vowel [aː] and the vowels posited in the middle except for [œː] ($p = .009$). For F2, significant difference was only found between [yː] and [uː] ($p < .001$). The development of vowel production by HSs from age 2;0 to 4;0 was inconsistent, with more expanded movement in F1 but not in F2. For example, F1 of [ɛː, iː, uː, œː, yː] decreased as age increased, resulting in the 4-year-old vowel space more contrastive in height. However, there was no significant change in F2 for all vowels; thus, the front and back vowels remained non-distinctive. The vowels produced by HSs at age 6;0 was observed to be more distributed compared to the vowel space at younger ages, and the relative position of each long vowel on the F1-F2 panel was similar to the reference. Statistical analysis further supported that [aː] had a significantly higher F1 frequency while [iː yː uː] had a significantly lower F1 than all the other vowels except for [œː] ([uː] vs [œː], p = .09), suggesting that HSs have acquired the vowel height of the primary vowels, although the space was still centralized relative to the reference pattern. Regarding vowel frontness, [iː, ɛː, yː] had a significantly higher F2 frequency while [uː] had a significantly lower F2 than all the other long vowels except for [ɔː]. Therefore, a vowel system with distinctive long vowels had been preliminarily formed at age 6;0 by HSs, with the high front vowels [iː, yː], a high back vowel [uː], a mid front vowel [ɛː], and a low central vowel [aː]. The vowel height of [œː] and the backness of [ɔː] were still developing. Furthermore, [iː] was expected to have a relatively lower F1 and higher F2 than [yː] and [ɛː]; however, there was a retraction of [iː] produced by HSs even at age 6;0, reflected by the overlap between [iː] and [yː] as shown in the F1-F2 panel.

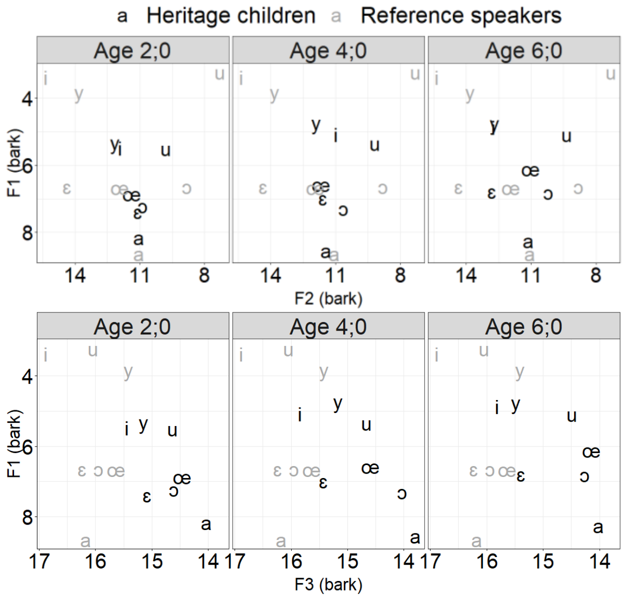

Figure 1: Mean F1, F2, and F3 values (in bark) of 7 long vowels produced by HSs and reference speakers.

3.2. Vowel roundedness contrast

Lip-rounding results in lowered F2 and F3 and is therefore useful in distinguishing between rounded and unrounded sounds. Cantonese has two pairs of vowels differing in lip roundedness: [iː] vs. [yː] and [ɛː] vs. [œː]. As seen in Figure 1, compared to the unrounded vowel [ɛː], the rounded vowel [œː] had a significantly lower F2 (p = .01) and F3 (p = .02) frequencies, revealing a significant distinction in lip-rounding between these two vowels by HSs at 6;0. The F3 distance between these two mid vowels was larger for HSs than for the reference speakers. Regarding the [iː yː] pair, HSs distinguished them in F3 with a lower frequency for [yː] than for [iː], although the difference was smaller than the reference. There was no significant difference between these two vowels in F2.

3.3. Vowel length contrast

Table 1 summarizes the durational differences between the mean value of [aː] and [ɐ] produced by the HSs, homeland children, and reference speakers in different environments. There was a significant effect for Age (F (3,915) = 21.254, $p < .001$), for Group (F (1,915) = 19.541, $p < .001$) and for Environment (F (1,915) = 888.094, $p < .001$). Evaluations regarding Age and Environment indicated that both HSs and homeland children utilized duration to distinguish [aː] and [ɐ] in pre-nasal environment since age 2;0. However, for the vowel contrast in diphthong, significant durational difference between these two vowels was only found in the homeland productions at age 6;0. There was no significant durational difference in each age group of HSs.

Table 1: The durational differences (in s) between the mean values of [aː] and [ɐ] by young HSs, homeland children, and reference speakers

	Diphthong	Pre-nasal
Heritage	0.08	0.17
Homeland	0.13	0.20
Reference	0.03	0.12

With respect to vowel quality, the formant data were analyzed using Gu's [26] Smoothing Spline ANOVA (SSANOVA). Simply put, any non-overlapping area represents significant difference between the formant measurements. Because of the page limit, only figures indicating the F1 and F2 trajectories of the youngest and the oldest target age groups were displayed here (Figure 2). The reference speakers produced [aː] and [ɐ] in diphthong [aːi ɐi] significantly differing in both F1 and F2, while HL and homeland children partially utilized the formant cue. HSs (Figure 2a) and homeland children (Figure 2c) were able to significantly distinguish the vowel length contrast since age 2;0, but the pattern of F2 was a reverse of those of the reference even at 6;0 (Figure 2b, 2d). For pre-nasal environments, while the reference speakers significantly produced the difference in both F1 and F2 between [aː] and [ɐ], HL and homeland children had a less contrastive pattern. HSs at age 2;0 significantly produced the difference in F1, but not in F2 (Figure 2e). Homeland children performed better than the HSs at age 2;0, with significant difference in both F1 and F2 (Figure 2g). However, no significant difference was found in F1 nor F2 among HSs (Figure 2f) and homeland children (Figure 2h) aged 4;0 and 6;0, suggesting that HL and homeland children could not distinguish [aː] and [ɐ] with pre-nasal coda in vowel quality.

To sum up, while reference speakers distinguish [aː] vs. [ɐ] in both vowel quantity and quality in all the environments, homeland children and HSs made use of the cues inconsistently according to environments.

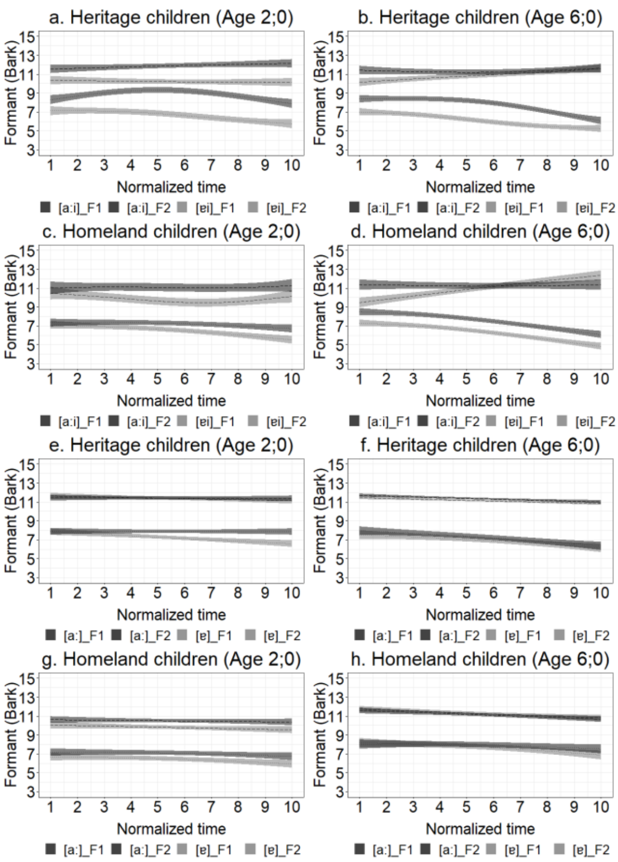

Figure 2: F1 and F2 trajectories for [aː] and [ɐ] in diphthongs (a-d) and before nasal codas (e-h) produced by HSs and homeland speakers.

4. DISCUSSION

This study investigated the acoustic development of 11 Cantonese vowels by young HSs. HSs mastered the distinction regarding vowel height earlier than vowel frontness. Unlike monolingual children who are reported to acquire the vowels by age 5;0, the acquisition of Cantonese vowels by HSs is not completed by age 6;0. For example, the vowels [œː], [ɔː], and [ɵ] were not mastered by 6-year-old HSs. It can be expected since even though the HSs had similar acquisition context as their monolingual peers at the early stage, they were exposed to the majority language since they attended schools at age 3;0. The interaction between two sound systems may influence the development of HL. For example, the absence of [œː] and [ɵ] in Toronto English vowel system reduced the amount of inputs of these vowels to the children, which may further slow their development down.

Regarding the vowel length contrast, HSs did not distinguish [aː] vs. [ɐ] in diphthongs in duration but partially in vowel quality. It can be expected since vowel quality functions as the primary cue in English tense/lax contrast [27]. For homeland children, they could distinguish [aː] vs. [ɐ] in both diphthongs and the pre-nasal environment by utilizing vowel quantity by age 6;0. Interesting pattern was found regarding the formant pattern of [aː] vs. [ɐ] in the pre-nasal environment. HSs did not have significant distinction in F1 and F2 since age 2;0, showing that they have not acquired this cue until age 6;0. However, for homeland children, significant differences in both F1 and F2 were found in their production at age 2;0, but the distinctions became weaker as they got older. Instances of [aː]-[ɐ] merger-in-progress have been found among native Hong Kong Cantonese speakers, in particular the younger generation [28], South Asian speakers [22], and Mandarin speakers [16]. A possible explanation is that, as age increased, the homeland children may communicate with a larger amount of people whose Cantonese may have such a merging phenomenon. Another interesting finding is with respect to the vowel roundedness contrast by HSs. While L2 learners found [ɛː] vs [œː] more difficult to distinguish than [iː] vs [yː] [22], HSs performed better in the distinction of the former pair. The extensively lower F3 for the rounded vowels relative to the unrounded vowel revealed that the HSs may overregularize the lip roundedness to discriminate the rounded vowel from the unrounded vowel. This finding provides evidence for the predictive model on the competence of HL sound system, supporting the principle that it is possible for the HSs to maintain or over-mark the phonologically distinct contrasts in HL. This study provided preliminary findings on the acquisition of vowels by young HSs. Data from more young HSs and from monolingual children at comparable ages will be adopted for a more comprehensive understanding.

5. ACKNOWLEDGEMENT

We would like to thank Professor Molly Babel and her RAs at the University of British Columbia for facilitating our data collection.

6. REFERENCES

[1] C. B. Chang, Y. Yao, E. F. Haynes, and R. Rhodes, Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustic Society of Ameirca*, *129*(6), 3964 – 3980, 2011.

[2] C. B. Chang and Y. Yao, Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal*, *13*(2), 134-160, 2016.

[3] W. M. Lam, Perception of lexical tones by homeland and heritage speakers of Cantonese. *University of British Columbia*. 2018.

[4] R. T. Y. Kan and M. S. Schmid, Development of tonal discrimination in young heritage speakers of Cantonese. *Journal of Phonetics*, *73*, 40-54, 2018.

[5] C. Lan and P. K. Mok, A preliminary study on Cantonese tone production by young heritage speakers. In *Proceedings of the 10th International Conference on Speech Prosody 2020*, 106-110. Tokyo, Japan. 2020.

[6] Statistics Canada. *Immigration and ethnocultural diversity highlight tables*. 2016 Census. Statistics Canada Catalogue no. 98-402-X2016007. Ottawa. 2017.

[7] Statistics Canada. *Canada [Country] and Vancouver, CY [Census subdivision], British Columbia (table). Census Profile.* 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. 2017.

[8] E. Zee. An acoustical analysis of the diphthongs in Cantonese. In *Proceedings of the XIVth International Congress of Phonetic Sciences*. 1999.

[9] E. Zee, Frequency Analysis of the Vowels in Cantonese from 50 Male and 50 Female Speakers. *Proc. 15th ICPhS*, Barcelona, 1117-1120. 2003.

[10] H.-N. S. Cheung. *A grammar of Cantonese as spoken in Hong Kong*. Hong Kong: The Chinese University of Hong Kong. 1972.

[11] O. Yue-Hashimoto. *Phonology of Cantonese*. Cambridge: Cambridge University Press. 1972.

[12] L. K. H. So and B. J. Dodd. The acquisition of phonology by Cantonese-speaking children. *Child Lang*. *22*, 473-495. 1995.

[13] S. F. Stokes and I. M. Wong. Vowel and diphthong development in Cantonese-speaking children. *Clinical Linguistics and Phonetics*, *16*, 597-617. 2002.

[14] C. To, P. Cheung and S. McLeod. A Population Study of Children's Acquisition of Hong Kong Cantonese Consonants, Vowels, and Tones. *Journal of speech, language, and hearing research: JSLHR, 56*. 2012.

[15] A. Holm and B. Dodd, A longitudinal study of phonological development of two Cantonese-English bilingual children. *Applied Psycholinguistics, 20*(3), 349-76. 1999.

[16] H. Tse, Beyond the monolingual core and out into the wild: A variationist study of early bilingualism and sound change in Toronto heritage Cantonese. Unpublished doctoral dissertation, University of Pittsburgh, 2019.

[17] M. Polinsky and G. Scontras. Understanding heritage languages. *Bilingualism: Language and Cognition 23*, 4–20. 2020.

[18] T. Kupisch. Towards modelling heritage speakers' sound systems. *Bilingualism: Language and Cognition 23*, 29–30. 2020.

[19] J. E. Flege. Second-language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp.233-278). Baltimore, MD: York Press. 1995.

[20] R. T. Y. Kan. Phonological Production in Young Speakers of Cantonese as a Heritage Language. *Language and Speech*, *64*(1), 73–97. 2021.

[21] A. Stoehr, T. Benders, J. G. Van Hell and P. Fikkert. Heritage language exposure impacts voice onset time of Dutch – German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. 2017.

[22] P. Mok, C.W. Leung, C. Lan and A.C.L. Yu. The acquisition of Cantonese vowel length contrast and vowel rounding contrast by South Asian students in Hong Kong. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*. Melbourne. 2019.

[23] P. Cheung, A. Ng, and C. K. S. To, *Hong Kong Cantonese Articulation Test*. City University of Hong Kong, Hong Kong, 2006.

[24] Y. Xu and H. Gao. FormantPro as a tool for speech analysis and segmentation. *Revista de Estudos da Linguagem*. 2018.

[25] H. Traunmüller. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, *88*(1), 97–100, 1990.

[26] C. Gu. *Smoothing spline ANOVA models*. New York: Springer New York, 2013.

[27] H. J. Giegerich. *English phonology: An introduction*. Cambridge, UK: Cambridge University Press. 1992.

[28] S. Y. R. Fung and C. K. C. Lee. Cantonese vowel merger-in-progress. In S. Calhoun, P. Escudero, M. Tabain , & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Australia (pp. 3225-3229). 2019.