JSLHR

Article

A Population Study of Children's Acquisition of Hong Kong Cantonese Consonants, Vowels, and Tones

Carol K. S. To,[a] Pamela S. P. Cheung,[b] and Sharynne McLeod[c]

Purpose: This study investigated children's acquisition of Hong Kong Cantonese.
Method: Participants were 1,726 children ages 2;4 to 12;4 (years;months). Single-word speech samples were collected to examine 4 measures: initial consonants, final consonants, vowels/diphthongs, and lexical tones. A 2-way analysis of variance was performed to examine the effects of age and sex on phoneme acquisition.
Results: There was rapid acquisition of initial consonants from age 2;6 to age 4;6. All 19 initial consonants were acquired by age 6;0 (90% criterion): /p-/, /m-/, and /j-/ were acquired the earliest; the last were /ts^h-/ and /s-/. Final consonants had a different acquisition time from their initial counterparts. Vowels were acquired by age 5;0 and diphthongs by age 4;0. All 9 tones were acquired by age 2;6. The main effect of age was significant for all 4 measures, whereas sex was significant for all measures except tone. Common phonological patterns (≥ 10%) for initial consonants were stopping, fronting, deaspiration, delabialization, affrication, and nasalization; patterns with 5.0%–9.9% occurrence were backing, deaffrication, gliding, and dentalization.
Conclusions: The acquisition of Cantonese showed similarities with English acquisition yet also had specific characteristics. Factors that contributed to the acquisition rate were functional load, articulatory ease, consonant–vowel interactions, phonetic variations, and the behavior of vowels and their allophones.

Key Words: speech sound development, normative study, consonant–vowel interactions, phonological patterns, Cantonese, tones, consonants, vowels, speech, children

Cantonese belongs to the Yue dialect, which is one of the seven dialect groups of modern Chinese. It is spoken by an estimated 55 million people throughout the world (Lewis, 2009), primarily in southern China, including Hong Kong, a special administrative region of China. According to the Hong Kong population by-census in 2006 (Census & Statistics Department, 2007), 90.8% of the local people age 5 years and older spoke Cantonese as the "usual language" and 5.7% spoke it as "another language/dialect" among a total population of 6.86 million in a small land area of 1,104 square kilometers.

[a]The University of Hong Kong, Hong Kong SAR, The People's Republic of China
[b]Child Assessment Service, Hong Kong SAR
[c]Charles Sturt University, Bathurst, New South Wales, Australia
Correspondence to Carol To: tokitsum@hku.hk

Editor: Anne Smith
Associate Editor: Barbara Davis
Received April 1, 2011
Revision received August 29, 2011
Accepted March 30, 2012
DOI: 10.1044/1092-4388(2012/11-0080)

Phoneme Acquisition and Phonological Patterns in Previous Cantonese Studies

In the past 2 decades, several cross-sectional studies on Cantonese speech sound acquisition have been conducted in Hong Kong. Previous comprehensive studies have used the approach of phoneme and phonological process analysis; for example, Cheung (1990) and So and Dodd (1995). More recent works have used nonlinear phonology and distinctive features to examine Cantonese sound development. For example, Wong and Stokes (2001) and Stokes and To (2002) examined feature acquisition of consonants, whereas Stokes and Wong (2002) investigated vowels and diphthongs. These studies have provided useful references regarding the acquisition of Cantonese features, phonemes, and developmental phonological patterns of typically developing children under the age of 6 years.

Cheung (1990) examined the speech samples of 155 typically developing preschool children ages 2;0 to 6;0 (years;months). They had not been diagnosed with speech sound disorders or had any reported hearing or learning difficulties. In the study, all initial consonants were acquired by 90% of the children by the age of 4;0 except /s-/, which was not acquired until age 5;6. Plosives, nasals, and approximants were acquired early, followed by fricatives and affricates. Initial front plosives were acquired before back plosives, and unaspirated sounds were acquired before their aspirated counterparts. Final consonants were acquired by age 4;0. Acquisition of tones, most vowels, and diphthongs occurred before age 2;0, except that /ɛ/ and /au/ were acquired by age 2;6, and /œ/ and /ai/ were acquired by age 3;0. Typical substitution processes included final consonant deletion, deaspiration of aspirated plosives/affricates, fronting of velars, stopping of fricatives/affricates (affrication of /s-/ → [ts-] was treated as stopping), delabialization of labiovelars (similar to cluster reduction), and lisping of /s-/. Cheung suggested that Cantonese-speaking children exhibited a similar order of sound acquisition and similar substitution processes as those of English-speaking children.

So and Dodd (1995) examined the speech of 268 typically developing children who were within the same age range and fulfilled the same exclusion criteria as the study participants in Cheung (1990). In the phoneme acquisition study of So and Dodd (1995), 90% of the children acquired all initial consonants by age 5;0 and all final consonants by age 4;6. Among the initial consonants, /s-/, /k^hw-/, and /ts^h-/ were the latest to be acquired, whereas boys and girls had different acquisition patterns for final consonants. Vowels, diphthongs, and tones were acquired by age 2;0. Phonological patterns that were used by more than 10% of the children were assimilation, /h-/ deletion, and cluster reduction of /kw-/ and /k^hw-/; stopping, fronting, deaspiration, and affrication of /s-/ to [ts-] and [ts^h-]; and deaffrication of /ts-/ and /ts^h-/ to [s-]. By age 4;0, less than 10% of children used these nonadult realizations consistently. The authors suggested that acquisition of Cantonese phonemes followed an order similar to that of English phonemes, although the rate of acquisition was faster. For phonological errors, assimilation, consonant cluster reduction, fronting, and stopping fitted well with the English patterns. However, fronting of /kw-/ to [p-], affrication of /s-/, and deaffrication of affricates were uncommon in English.

For vowel and diphthong development, Stokes and Wong (2002) suggested that both ambient frequency and feature complexity might contribute to the development of vowels and diphthongs in Cantonese. The ranked order of vowel accuracy at 24 to 27 months, from easy to difficult, was ɛ > ɔ > a = i > ɐ > œ/ɵ > u > y,[1] whereas that of diphthong accuracy at the same age range was ou > ɐi > ei > œi > iu > ɔi > ai = au > ɵy > ui.

Effect of Age and Sex on Speech Sound Acquisition

Age

Whereas English-speaking children may not have acquired speech sounds until ages 7;0 to 9;0 at the 90% acquisition criterion (Dodd, Holm, Zhu, & Crosbie, 2003; Smit, Hand, Freilinger, Bernthal, & Bird, 1990), Cantonese-speaking children have acquired their sound system by age 5;0–5;6 (Cheung, 1990; So & Dodd, 1995) using the same criterion. The smaller number of consonants and the lack of consonant clusters in Cantonese may lead to this early acquisition.

Sex

English studies report that there is a sex effect, favoring girls, in speech sound acquisition (e.g., Dodd et al., 2003; Kenny & Prather, 1986; Smit et al., 1990). In a meta-analysis of sex differences in verbal ability (Hyde & Linn, 1988), it was found that, although there was a small overall effect size for sex in 165 studies, speech production had the largest effect size for female superiority among other verbal abilities like vocabulary or reading comprehension. No sex effect was found in the two Cantonese studies. So and Dodd (1995) did not report inferential statistics on sex effect, but descriptive statistics showed that the girls had a more rapid acquisition of initial consonant between the ages of 2;0 and 3;9. Cheung (1990) did not demonstrate any significant sex effect in any age group either, although it was also observed that the girls were faster than the boys in the acquisition of initial consonants between the ages of 2;0 and 3;6. One reason why the sex effect was not significant in Cheung (1990) might be attributed to the low statistical power resulting from the small group size, as each of the 16 groups had an average of only 10 participants. A larger group size may help reveal any sex effect.

Components of Cantonese

In the present study, we adopted the International Phonetic Alphabet system, which was used in the phonetic analysis of Hong Kong Cantonese by Zee (1999). Cantonese has the syllabic structure of (C)V/C_{syll}(C). That is, a syllable may have the form CV, VC, CVC, V, or syllabic consonant (C_{syll}). A syllabic consonant is a word that does not contain a vowel; for example, /ŋ$_5$/ (*five*). Most Cantonese words are disyllabic; however, words also can be monosyllabic, trisyllabic, and, very occasionally, polysyllabic.

[1]The symbol ">" indicates that the phoneme on the left is acquired earlier than the one on the right; the symbol "=" indicates that the two phonemes are acquired at a similar time.

Consonants

There are 19 initial consonants and six final consonants in standard Cantonese. For initial consonants, the five pairs of voiceless aspirated and unaspirated plosives/affricates are /p-, p^h-, t-, t^h-, k-, k^h-, k^w-, k^{wh}-, ts-, ts^h-/.[2] There are three fricatives, /f-, s-, h-/; two approximants, /j-, w-/; and one lateral, /l-/. The three nasals, /m, n, ŋ/, appear as both initial and final consonants. Among them, /m/ and /ŋ/ can behave as syllabic segments. Three plosives—namely, /-p, -t/, -k/—also act as final consonants and are unreleased.

Vowels and Diphthongs

There are 11 vowels and 11 diphthongs in Cantonese (Zee, 1999). Seven of them are long vowels, /i, y, ɛ, œ, a, ɔ, u/, and four are short vowels /ɐ, ɪ, ɵ, ʊ/. The short vowels /ɪ, ɵ, ʊ/ are allophones of /i, œ, u/, respectively. The 11 diphthongs are /ai, ei, ɐi, ui, ɔi, au, ɐu, iu, ou, ɵy, ɛu/.

Tones

Cantonese is a tone language that uses pitch difference, in terms of fundamental frequency, to indicate different lexical meanings. There are nine tones in Cantonese, and they are numbered from T1 to T9 in this article (see Figure 1). Syllables with T1 to T6 end with a vowel, a diphthong, or a nasal, whereas syllables with T7 to T9 end with a plosive /-p, -t/ or /-k/. The tone contour of T1 and T7 is high, that of T2 is mid-low to high, that of T3 and T8 is mid-range, that of T4 is mid-low to low, that of T5 is mid-low to mid-range, and that of T6 and T9 is mid-low. T7, T8, and T9 share the same tone contour as T1, T3, and T6, respectively, except that they are shorter in duration. The tone digits in Figure 1 represent the relative tone values of these nine tones. In a transcribed syllable, the tone number, if indicated, is marked on the right of the transcription.

Syllable Positions for Target Consonants

In English, a consonant in word-initial, within-word, and word-final positions may be acquired at different rates (e.g., Smit et al., 1990); therefore, a target consonant is tested in different word positions, including word initial, word medial, and word final. Whereas English is spoken with a stress-timed rhythm, Cantonese is spoken with a syllable-timed rhythm. Each syllable in a word receives approximately equal emphasis and generally takes an equal interval of time (Bauer & Benedict, 1997). With this property, syllables are often considered to be the basic independent pronunciation units for Cantonese. Each syllable can be represented by a Chinese character, which is the smallest nonseparable lexical unit that can be pronounced. The complexity of producing the same consonant at word-initial or within-word position in Cantonese is similar. In other words, the within-word position in Cantonese consonant acquisition may not be as important as in English. Studying the development of consonants in the syllable-initial and syllable-final positions is considered to be adequate for Cantonese, and examining consonants in these two positions has been used by previous studies on developmental phonology in Cantonese (Cheung, 1990; So & Dodd, 1995; Stokes & To, 2002).

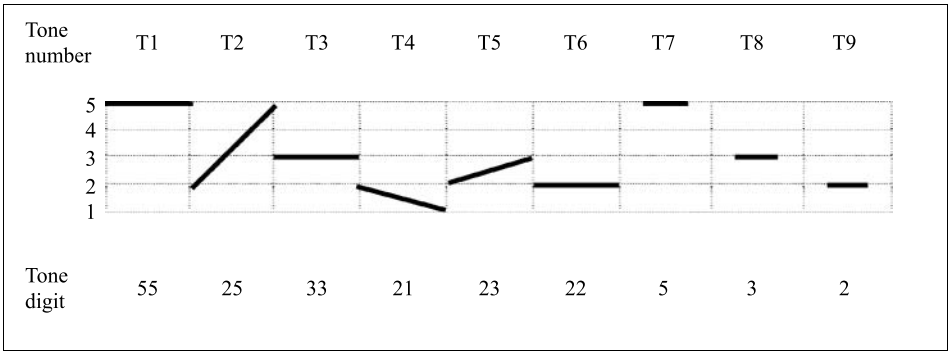
Phonetic Variations and Ongoing Sound Changes in Hong Kong Cantonese

When examining children's speech acquisition, it is important to define what are acceptable and unacceptable productions compared with the ambient language. For example, Smit et al. (1990) provided an appendix containing acceptable adultlike productions of phonemes for the Iowa–Nebraska Articulation Norms Project. In the past few decades, ongoing sound changes in Hong Kong Cantonese have been noted and investigated. The present study accepted nine phonetic variants that are commonly used by adult speakers of contemporary Hong Kong Cantonese (see Appendix). The first five variants were also considered in Cheung (1990) but were not mentioned in So and Dodd (1995) or So (1993) when examining Cantonese children's speech acquisition.

These variants were identified in To, Cheung, and McLeod (2012) through a historical analysis of documents and a comparison between standard pronunciations as documented in dictionaries with productions of single words by 112 native Cantonese-speaking adults (18–45 years old) in Hong Kong. Among the nine phonetic variants that occur in contemporary Hong Kong Cantonese, five were related to initial and syllabic consonants: (1) /n-/ → [l-], (2) /ŋ-/ → Ø-, (3) Ø- → [ŋ-], (4) /k^wɔ-/ → [kɔ-], and (5) syllabic /ŋ/ → [m]. These five variants were examined in one stimulus word each. The rest were related to final consonants: (6) /-ŋ/ → [-n], (7) /-n/ → [-ŋ], (8) /-k/ → [-t], and (9) /-t/ → [-k]. Varying numbers of stimulus words were examined. Phonetic variation (1), /n-/ → [l-], nearly merged with the standard form, as there was a 94.6% occurrence of [l-] among the adults studied. Phonetic variations (2), /ŋ-/ → Ø-, (3), Ø- → [ŋ-], and (4), /k^wɔ-/ → [kɔ-], were commonly produced by the adults, with an occurrence of 31.3%, 37.5%, and 50.9%, respectively. Phonetic variation (5), syllabic /ŋ/ → [m], almost replaced the standard form, with an occurrence of 95.5% in syllabic [m]. Phonetic variations (6) through (9), which related to syllable-finals, were vowel dependent and occurred less frequently than did variations (1) through (5). Phonetic variation (6), /-ŋ/ → [-n], was examined in six vowel

[2]Some authors (e.g., So & Dodd, 1995) have classified /k^w-/ and /k^{wh}-/ as consonant clusters /kw-/ and /k^hw-/; in the present study, these are classified as consonants according to the conventions of Zee (1999).

Figure 1. Nine Cantonese tones. T = tone.



contexts—namely, /ɪ, ɔ, ɛ, ɐ, a, œ/[3]—and had an average occurrence of 24.1% among the adults. In further analysis, the change varied from 0% in /ɪŋ/ to 37.5% in /œŋ/. Four vowel contexts—namely, /ɛ, ɐ, a, œ/—were more likely to result in the production of [-n]. Phonetic variation (7), /-n/ → [-ŋ], which was a reverse pattern from variation (6), was examined in seven vowel contexts—namely, /i, y, u, ɵ, ɐ, a, ɔ/. It had an average occurrence of only 1.1% and was most frequently found in the /ɔn/ vowel context (7.1%). Phonetic variations (8), /-k/ → [-t], and (9), /-t/→[-k], were examined in the vowel contexts of /ɪ, ʊ, œ, ɔ/ and /i, a, ɐ, ɔ/, respectively. Phonetic variation (8) had an average occurrence of 6.0%. It was mostly produced in the rimes of /ɔk/ and /œk/, 16.1% and 8.0%, respectively, but 0% in /ɪk/ and /ʊk/. Phonetic variation (9), /-t/ → [-k], had an average occurrence of 6.3%, with 0% in the vowel contexts of /it/ and /at/ but 24.1% in /ɔt/ and 0.9% in /ɐt/.

Interactions Between Consonants and Vowels

CV interactions have been noted in the speech of young typically developing children across many languages. Davis and MacNeilage (1995) studied the babbling patterns of six infants over a 4- to 6-month period and observed CV interdependence: Front vowels would preferentially co-occur with alveolar consonants, back vowels with velar consonants, and central vowels with labial consonants. The authors proposed the "frames, then content" hypothesis for speech acquisition, in which the syllabic "frame" of rhythmic mandible oscillation produces predicted babbling patterns and within which segment-specific "content" elements develop. This CV preference was observed not only in babbling but also during the phonological development in early speech. For example, Davis, MacNeilage, and Matyear (2002) longitudinally studied the early word production of 10 infants and found additional support for the "performance-related frame/content approach" compared with the "competence-based approaches . . . within the generative phonological tradition" (p. 75).Vihman (1992) studied the first 50 words of 23 children across four languages—namely, American English, French, Japanese, and Swedish—and showed that the co-occurrences of the sequence of bilabial consonants with central vowels and velar consonants with back vowels were strongly supported in all four languages. Although the alveolar consonants and front vowels association was observed, the association was not as strong as that between the former two. Sussman, Duder, Dalston, and Cacciatore (1999) reported a longitudinal case study of a monolingual English-speaking girl from age 7 months to age 40 months by analyzing the second formant (F2) onset and F2 vocalic centers of her English productions. Alveolar stops yielded the most fronted vowels, initial velar stops yielded the most back vowels, and initial labials yielded the most centered vowels. Clinical data on Cantonese-speaking children with speech sound disorders also provide support for CV interaction. Analysis of disordered speech in Cantonese shows that some children who had difficulties producing the alveolar/palatal phonemes /t-, t^h-, ts-, ts^h-, j-/ could achieve the target sounds if they were followed by the high-front vowels /i/ or /y/ (Cheung & Abberton, 2000). This implies that a vowel with a place of articulation similar to the initial consonant may facilitate the production of that consonant, or that vowels with a distant position from the initial consonant may, on the contrary, hinder the production. This clinical evidence suggests that children with typical speech development may acquire a consonant with different vowel combinations at different rates.

Given the evidence of CV interaction during early speech acquisition, the stimuli for consonants selected in the present study included items with various vowel combinations. CV interaction in the design of assessment materials has not been considered in previous phonological studies of English or Cantonese. If this design was used in English, the number of test items would increase

[3]The rime /ʊŋ/ was examined in a child group (age 11;0), but not in the adult group, and the occurrence of nonstandard production was 0%.

considerably. However, it is feasible to consider this factor in Cantonese, as there are fewer initial consonants in Cantonese than in English (19 compared with 23), fewer final consonants (six compared with 21), and no Cantonese consonant clusters.

The Present Study

The present study sought to examine speech sound acquisition of a large number of Cantonese children. There were four main reasons why a large-scale study of Cantonese speech acquisition was necessary: (a) The effect of age and sex had not been fully explored; (b) there were no large-scale studies of Cantonese children's production of consonants, vowels/diphthongs, and tones that systematically described typical nonadult realizations[4] (cf. Smit, 1993) and phonological patterns; (c) current studies of Cantonese speech acquisition have not taken into account the phonetic variations that have resulted from the ongoing sound changes in Hong Kong, so phonetic variants were previously counted as incorrect; and (d) to date, no Cantonese speech acquisition study has considered the interaction between consonants and vowels. To address the aforementioned issues, a larger scale study with a more representative sample and more carefully selected stimuli was conducted to examine the speech sound development in Cantonese. With the stronger research design, the present study aimed to (a) explore the possible effect of age and sex, (b) compare the findings from the present study with previous acquisition studies in Cantonese and English, and (c) evaluate the contribution of CV interaction in the acquisition of consonants.

Method

Participants

To obtain an objective picture of the acquisition of Cantonese phonology and to allow adequate statistical power to detect potential age and sex effects, the required sample size was estimated using the sample size calculator G*Power (www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/; Erdfelder, Faul, & Buchner, 1996). To achieve a confidence level of 0.95 and a power level of 0.8 with an expected effect size of 0.1, the ideal number of participants required to detect the potential age and sex effect was 1,797, and approximately 128 participants (50% male and 50% female) were needed for each age group. A stratified random sampling method was adopted with three stratifying variables of age, sex, and geographic district. Children were recruited from preschools and primary schools in three Hong Kong districts (Hong Kong Island, Kowloon, and the New Territories) with different socioeconomic status as measured by median monthly income based on the 2006 Census information (Census & Statistics Department, 2007). In each school, the class lists were used as the sampling frames. The final sample consisted of 1,726 participants. The group size of each age and sex ranged from 50 to 80 participants, with an average of 65.64 (see Table 1).

A total of 852 boys and 874 girls who spoke Cantonese as their major language were recruited. The children spanned in age from 2;4 to 12;4; that is, from preschool children to students who were in their last year of primary school education. Because the sound system of a child expands rapidly in early childhood, children between ages 2;4 and 5;7 were grouped into eight 6-month age bands (except for the age 5;6 group, which had a 4-month age range because of primary school entry at age 5;8), and primary school children from age 6;0 to 11;0 were divided into six 1-year age bands (see Table 1). With the participation of primary school students, mastery of phonemes could be shown without age restriction. Any nonadult realizations of older children could also be revealed.

Many nursery schools and primary schools in Hong Kong provide inclusive education and take a small number of children with intellectual or other disabilities. Previous cross-sectional Cantonese studies excluded participants with reported problems of intelligence, hearing, speech, or language (Cheung, 1990; So & Dodd, 1995). There has been controversy around whether to include or exclude children with language problems when developing a normative reference for speech and language development. Peña, Spaulding, and Plante (2006) recommended the exclusion of children with impaired language skills from the normative group if the purpose of testing was to identify children with language impairment. The authors demonstrated that the sensitivity of a test would be lowered if a mixed normative group that included a clinical sample was used instead of a pure normative group that included only the typical sample. The only tradeoff of a pure normative group was that false positives would increase with the lowering of the specificity. However, Peña et al. (2006) added that a broad representative sample of children of all capabilities was appropriate if the interpretation of the test scores was relative rather than absolute—for example, to document the degree of difference from the mean of the general population rather than that of the typical population. On the other hand, McFadden (1996) commented that a sample with exclusion criteria did not represent the full range of the normative sample, as children with special needs would presumably score at the lower end of a general population. Dodd et al. (2003) also agreed that including all children in the normative data avoided overidentification

[4]The terms *developmental error*, *common mismatch*, and *nonadult realization* have been used to describe children's production of phonemes, each with a slightly different meaning. In the present article, we have chosen the neutral term *nonadult realization.*

Table 1. Age and sex of the participants.

Age group	Range (years;months)	*SD* (years;months)	Mean age		*n*		*N*
			Male	Female	Male	Female	
2;6	2;4–2;9	0.15	2.56	2.56	54	50	104
3;0	2;10–3;3	0.13	3.05	3.08	67	64	131
3;6	3;4–3;9	0.13	3.51	3.53	58	63	121
4;0	3;10–4;3	0.10	4.06	4.02	59	54	113
4;6	4;4–4;9	0.14	4.49	4.47	60	80	140
5;0	4;10–5;3	0.10	5.04	5.02	66	66	132
5;6	5;4–5;7	0.15	5.56	5.52	69	68	137
6;0	5;8–6;7	0.19	6.27	6.24	66	72	138
7;0	6;8–7;7	0.19	7.21	7.25	67	75	142
8;0	7;8–8;7	0.21	8.27	8.20	72	71	143
9;0	8;8–9;7	0.23	9.22	9.23	73	73	146
10;0	9;8–10;7	0.26	10.26	10.26	71	70	141
11;0	10;8–12;4	0.29	11.26	11.25	70	68	138
Total					852	874	1,726

of typically developing children who fell at the lower end of a normal distribution curve. Because the aims of the present study were to examine speech sound development and the phonological patterns of the general population, inclusion of all children provided a complete view of sound acquisition and did not truncate nonadult realizations in general development. Therefore, participants with reported or observed difficulties in intelligence, hearing, speech, or language were not excluded, and the sample represented the population of Hong Kong as a whole.[5]

Assessors and Transcribers

The assessors were 11 qualified and experienced pediatric speech-language pathologists (SLPs). They received training on the test procedures to administer the test in a standardized way. Two SLPs with more than 10 years of experience in assessing child speech and language were responsible for transcribing and scoring the speech samples following standardized procedures.

Speech Sampling Tool

A single-word sampling tool, the beta version of the Hong Kong Cantonese Articulation Test (HKCAT; Cheung, Ng, & To, 2006), was used in the present investigation. A total of 42 color photographs were used to elicit 51 familiar words containing all the Cantonese consonants, vowels, diphthongs, and tones. Certain objects were used to elicit two target words, which could be either an object label or a physical description of the object. The 19 initial consonants and six final consonants were elicited with the consideration of the CV interaction effect. That is, each consonant was preceded or followed by a front, central, or back vowel. One stimulus word in each sound environment was chosen as the scoring item. In most cases, there were three scoring items for each consonant. Exceptions were found in /k^{w}-, k^{wh}-, n-, w-, ŋ-/ and in the null onset Ø- as restricted by Cantonese phonotactic rules or rare syllables that were not picturable. This yielded a total of 48 stimuli for the 19 initial consonants and 16 stimuli for the six final consonants.

The seven long vowels appear in both open and closed syllables and thus were sampled in both positions. The short vowels appear only in closed syllables, and one stimulus word was chosen for each. Diphthongs appear only in open syllables, and one stimulus word was chosen for each as well. A total of 18 and 11 stimuli were selected for the scoring of vowels and diphthongs, respectively. For the nine lexical tones, two stimuli for each were selected, resulting in 18 stimuli. To summarize, elicitation of the 51 words enabled consideration of 48 productions of initial consonants, 16 productions of final consonants, 18 productions of vowels, 11 productions of diphthongs, and 18 productions of tones.

Procedure

The participants were tested individually in a quiet room in their nursery schools, kindergartens, or schools. When a rapport was established, the stimulus book was presented. A standard carrier question was asked to elicit each target word. If the participant could not label the picture, the assessor provided the target word for imitation. Imitated responses occurred infrequently, mostly produced by the very young children. An imitated response was treated in the same way as a spontaneous response in

[5]According to the parent questionnaire in the present study, parents reported that 2.3% of the participants had a history of receiving speech and language therapy, but the nature of the therapy—whether speech or language—was not specified.

the scoring. Participants were encouraged to repeat a certain production if it was not spoken clearly for quality recording. Speech samples were recorded by mini-disc (MD) recorders (either Sony Mz-B100 portable MD recorder or Sharp MD-MT290H(S) MD recorder). A lapel microphone (either the ECM-DS70P Sony electret condenser microphone or the RP-VC 200 Panasonic stereo microphone) was attached to each participant's clothing approximately l5 cm from the mouth. After the testing, the two experienced assessors listened to all the recordings and transcribed and scored the samples.

Analysis

Overall measures. To compare the performance among age groups, between sexes, and the interaction of these two factors, the raw scores for four measures—namely, initial consonants, final consonants, vowels/diphthongs, and tones—for each participant were calculated as the outcome measures. These four measures refer to the percentage of all the selected items with correct production. The present study also considered the nine completed or ongoing sound changes mentioned previously as acceptable variants.

Effect of age and sex. A two-way Age (13 age groups) × Sex (2 sexes) analysis of variance (ANOVA) was conducted for each of the four overall measures: initial consonants, final consonants, vowels/diphthongs, and tones. Post hoc trend analyses were performed to examine the linear relationship of the age group effect.

Age of phoneme acquisition. A phoneme was considered to be acquired if it was produced correctly in all sound environments. In the present study, an initial or final consonant needed to be produced correctly in all three vowel contexts (excluding the exceptions as discussed previously) to establish that it had been acquired. Similarly, long vowels needed to be produced correctly in both open and closed syllables, short vowels in closed syllables, and diphthongs in open syllables. Tones were considered to be acquired if the two scoring items for each tone were correct.

Age of acquisition was documented for 90% of children. The 90% criterion demonstrates the age of mastery of a particular phoneme and has been widely used in English and Cantonese studies on developmental phonology (Cheung, 1990; Dodd et al., 2003; Smit et al., 1990; So & Dodd, 1995). The data were compared with previous Cantonese and English studies that used this criterion.

Nonadult realizations and phonological patterns. The percentage of the nonadult realizations over the total productions of the scoring items for each target phoneme in each age band was calculated. Nonadult realizations that made up 5.0%–9.9% and ≥ 10% of the total productions of a particular phoneme were listed. Similarly, phonological patterns that had 5.0%–9.9% and ≥ 10% occurrence were summarized for initial consonants, final consonants, and vowels/diphthongs.

CV interactions. To study CV interactions in consonant acquisition, we examined the phonological patterns of consonants in relation to adjacent front, central, and back vowels. We observed the degree of vowel effect by comparing the percentages for a certain type of substitution of a phoneme when attached to different vowels. To look for any consistent patterns related to vowel context and to minimize the effects of other factors, we investigated realizations that differed from the target phoneme by only one process. For example, substitutions involving the single affrication process (e.g., /t-/ → [ts-]) were examined, but those involving both affrication and aspiration (e.g., /t-/ → [ts^h-]) were not included. In this way, any difference in the percentages of nonadult realizations for the consonant when combined with a front, central, or back vowel could be claimed as coming mainly from vowel effect, although the influence caused by the other consonant in the syllable cannot be totally eliminated. Only initial alveolars and velars were included in the CV interaction analysis. Specifically, we examined backing of /t-/ → [k-] and /t^h-/ → [k^h-]; fronting of /k-/ → [t-] and /k^h-/ → [t^h-]; stopping of /ts-/ → [t-], /ts^h-/ → [t^h-], and /s-/ → [t-]; affrication of /t-/ → [ts-], /t^h-/ → [ts^h-], and /s-/ → [ts-]; and deaffrication of /ts-/ → [s-] and /ts^h-/ → [s-]. Bilabials were not studied because /p-/ and /m-/ were acquired early and did not have clear error patterns for analysis. CV interaction in final consonants was not examined, as there were phonetic variants in the production of /-n, -ŋ, -t, -k/. Only the speech samples of the age groups from age 2;6 to age 6;0 were used because all Cantonese initial consonants were acquired by 90% of the children by age 6;0.

Interrater Reliability

The interrater reliability of the two transcription assessors was examined. A total of 130 of the 1,726 speech samples (7.5%; 5 males and 5 females from each age band) were randomly selected and were independently scored by the other assessor. Point-to-point agreement between the two assessors was 98.3% for the 31,850 phonemes and tones of both the scoring and nonscoring items.

Results

General Trends in Speech Acquisition

The mean percentages and standard deviations for the accuracy scores of initial and final consonants, vowels/diphthongs, and tones of the 13 age groups, by sex, are summarized in Table 2. Sound production accuracy improved with age for all four measures. For initial consonants, there was a rapid rise in accuracy from age 2;6 to

Table 2. Mean percentage (and *SD*) of accuracy by age group and by sex on the four overall measures.

		Initial consonant		Final consonant		Vowel/diphthong		Tone	
Age group and sex	***n***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
2;6									
Female	50	78.12	16.31	87.63	17.79	90.83	8.29	98.89	2.75
Male	54	69.73	19.54	77.43	23.34	89.72	11.94	97.22	6.63
Total	104	73.76	18.46	82.53	0.90	90.25	10.31	98.02	5.19
3;0									
Female	64	84.57	15.60	89.26	17.20	94.34	7.15	99.48	1.63
Male	67	79.41	16.88	87.97	14.24	94.60	7.06	99.09	2.84
Total	131	81.93	16.41	88.61	0.81	94.47	7.08	99.28	2.33
3;6									
Female	63	90.86	10.37	92.56	9.25	95.18	4.53	99.21	1.96
Male	58	90.71	11.89	92.78	11.99	96.61	6.34	99.52	1.57
Total	121	90.79	11.08	92.67	0.84	95.87	5.50	99.36	1.78
4;0									
Female	54	93.42	9.34	94.91	11.26	97.06	3.69	99.59	1.47
Male	59	93.57	8.06	88.35	18.69	96.14	8.25	99.15	2.01
Total	113	93.50	8.66	91.63	0.87	96.58	6.48	99.36	1.78
4;6									
Female	80	97.47	3.28	97.34	6.19	98.32	2.74	99.24	2.30
Male	60	94.32	8.64	95.42	8.83	98.28	3.01	99.35	1.80
Total	140	96.12	6.35	96.38	0.79	98.30	2.85	99.29	2.09
5;0									
Female	66	97.34	4.78	95.93	10.31	97.54	4.18	100.00	0.00
Male	66	96.32	7.15	96.88	5.62	98.54	2.35	99.75	1.17
Total	132	96.83	6.08	96.40	0.80	98.04	3.41	99.87	0.83
5;6									
Female	68	97.90	3.75	98.35	3.17	99.14	1.92	99.18	2.20
Male	69	96.21	6.69	96.47	6.56	98.35	2.68	99.36	2.03
Total	137	97.05	5.48	97.41	0.79	98.74	2.36	99.27	2.11
6;0									
Female	72	98.33	2.48	98.70	3.31	98.66	1.97	99.38	1.76
Male	66	98.18	3.51	96.59	9.82	98.38	3.10	98.91	2.79
Total	138	98.25	3.01	97.64	0.79	98.53	2.56	99.15	2.32
7;0									
Female	75	99.13	1.65	98.92	2.59	99.13	1.89	99.63	1.40
Male	67	98.08	3.41	97.29	5.01	98.61	2.00	99.09	2.07
Total	142	98.63	2.67	98.11	0.77	98.88	1.95	99.37	1.76
8;0									
Female	71	99.28	1.47	98.33	3.65	99.27	2.09	98.98	2.71
Male	72	99.18	1.74	97.74	4.24	99.19	1.78	99.54	1.55
Total	143	99.23	1.60	98.04	0.77	99.23	1.94	99.26	2.21
9;0									
Female	73	99.61	1.00	99.23	2.75	99.62	1.36	99.32	1.84
Male	73	99.30	1.33	98.80	2.48	99.39	1.45	99.54	1.54
Total	146	99.45	1.18	99.02	0.76	99.50	1.40	99.43	1.69
10;0									
Female	70	99.01	1.58	98.66	3.35	99.01	1.96	98.89	2.61
Male	71	98.99	1.92	98.86	3.05	99.32	1.90	98.98	2.54
Total	141	99.00	1.75	98.76	0.78	99.17	1.93	98.94	2.56
11;0									
Female	68	99.37	1.23	98.16	5.41	99.14	1.62	99.02	2.53
Male	70	99.01	2.59	99.20	3.67	99.16	2.84	98.81	3.40
Total	138	99.19	2.04	98.68	0.78	99.15	2.31	98.91	2.99

Note. Ages are given as years;months.

age 4;6, which slowed down and reached a plateau from 6 years onward. For final consonants, approximately 90% of correct productions had been achieved by age 3;6. Vowels and diphthongs were acquired early, with 90% of correct productions achieved by age 2;6, approaching completion by age 4;6. The lexical tones were acquired the earliest among the four domains and had reached an average of 98% acquisition at age 2;6.

Effects of Age and Sex in Speech Sound Development

Girls generally performed better than boys in most age groups, particularly in the younger ages (see Table 2). This observation generally was supported by the ANOVA results. For the measures of initial consonants, final consonants, and vowels/diphthongs, the main effects of age group and sex were significant: For initial consonants, age group, $F(12, 1699) = 141.89$, $p < .001$, partial $\eta^2 = .50$; sex, $F(1, 1699) = 30.79$, $p < .001$, partial $\eta^2 = .02$. For final consonants, age group, $F(12, 1700) = 31.57$, $p < .001$, partial $\eta^2 = .18$; sex, $F(1, 1700) = 16.33$, $p < .001$, partial $\eta^2 = .01$. For vowels/diphthongs, age group, $F(12, 1699) = 34.73$, $p < .001$, partial $\eta^2 = .20$; sex, $F(1, 1699) = 10.31$, $p = .001$, partial $\eta^2 = .01$. Trend analyses of age group for the three measures reveal that there was also a linear component to the relationships (all had ps $< .001$), with higher average performance for the older groups than for the younger ones. For the measure of tones, the main effect of age group, $F(1, 12) = 3.35$, $p < .001$; partial $\eta^2 = .02$, but not sex, was significant, $F(1, 1699) = 1.41$, $p = .25$, partial $\eta^2 = .001$. The main effect of age group had the larger effect size (as indicated by the partial η^2 value), accounting for more variance for all four measures (50.1% of the scores for initial consonants, 11.5% for final consonants, 19.7% for vowels/diphthongs, and 2.3% for tone) than that of sex.

A significant Sex × Age Group interaction effect was found only in initial consonants, $F(12, 1699) = 3.43$, $p < .001$, partial $\eta^2 = .02$, and final consonants, $F(12, 1700) = 3.12$, $p < .001$, partial $\eta^2 = .02$. The different sex effects in all age groups in initial-consonant and final-consonant acquisition are shown in follow-up plots of age group by percentage accuracy (see Figure 2).

Age of Phoneme Acquisition

Tables 3 and 4 show the age of phonemic acquisition of initial consonants, final consonants, and vowels/diphthongs by at least 90% of participants of each age band, respectively. The available results in the studies by Cheung (1990) and So and Dodd (1995) are included for comparison at the 90% acquisition criterion.

Figure 2. Sex effect in different age groups in (a) initial consonant and (b) final consonant acquisition.

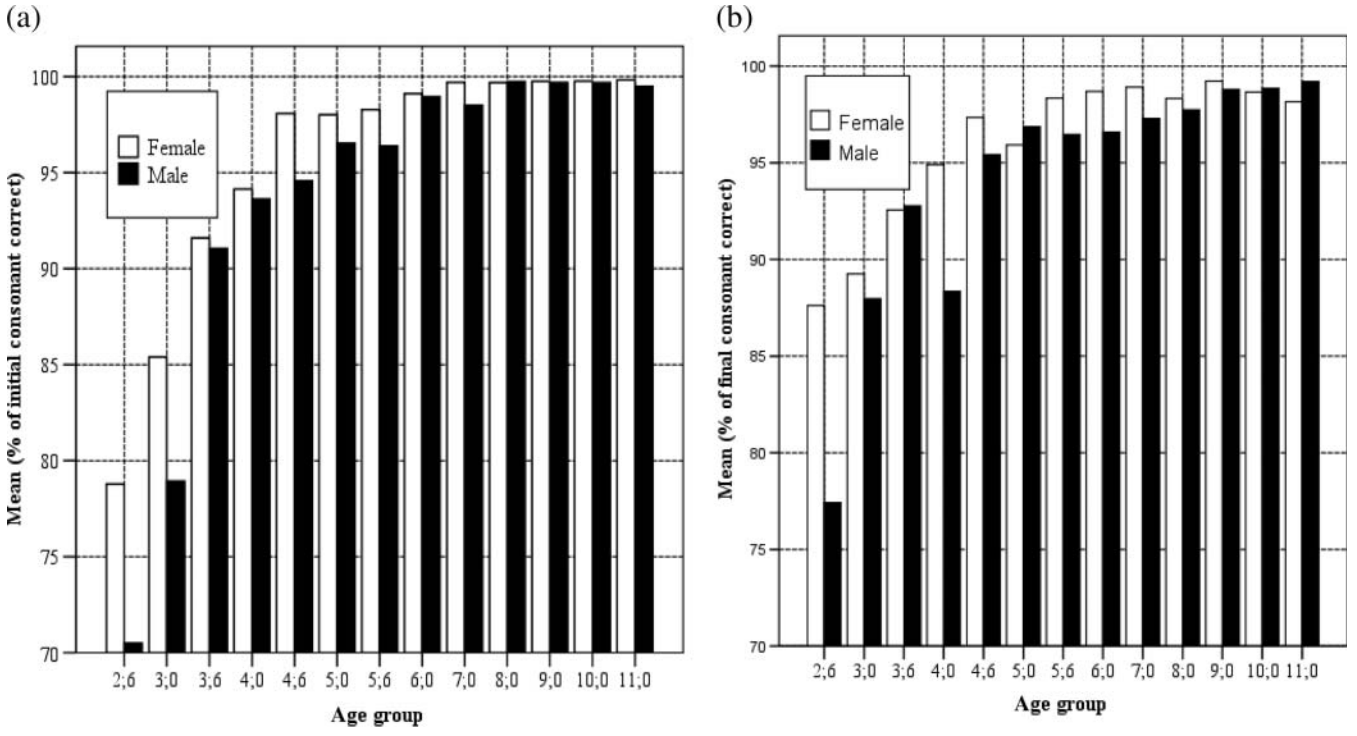


Table 3. Age of acquisition of initial consonants in the present study compared with studies by So and Dodd (1995) and Cheung (1990).

Age group	Present study		So & Dodd (1995)[a]		Cheung (1990)[b]
	Male	Female	Male	Female	All
2;6	p-	p-, ŋ-/Ø-	(n-, ŋ-: by 2;0), p-, t-, j-	(p-, m-, n-, h-, w-, j-: by 2;0) (t-: by 2;3) k-	(p-, t-, m-, j-: by 2;0) n-/[l-], l-, h-, w-
3;0	j-, m-	j-, w-, m-	m-, w-	(ŋ-, l-: by 2;9) f-	ŋ-/Ø-, p^h-, t^h-
3;6	p^h-, f-, h-, w-, ŋ-/Ø-	p^h-, t-, f-, h-	(k-: by 3;3) h-	(p^h-: by 3;3) k^h-	k-, k^w-, k^{wh}-
4;0	t-, l-	k-, k^h-, k^w-, k^{wh}-, l-	p^h-, t^h-, k^h-, ts-, kw-, l-	(kw-: by 3;9) t^h-	f-, ts-, ts^h-, k^h-
4;6	t^h-, k^w-, ts-	t^h-, ts-, ts^h-	(f-, ts^h-: by 4;3) s-	(ts-: by 4;3) s-	
5;0	k-, k^h-, k^{wh}-		k^hw-	(k^hw-: by 4;9) ts^h-	
5;6					s-
6;0	ts^h-, s-	s-			

Note. Ø = null realization.

[a]So and Dodd (1995) classified /k^w-/ and /k^{wh}-/ as consonant clusters /kw-/ and /k^hw-/, respectively. [b]The phonetic symbols used in Cheung (1990) were translated to the International Phonetic Alphabet as used in the present study.

Initial consonants. Ninety percent of the children acquired all Cantonese initial consonants by age 6;0. In terms of place and manner of articulation, bilabials were acquired early, followed by alveolars, velars, and then labiovelars; plosives were acquired before fricatives and affricates. In general, girls acquired the initial phonemes slightly faster than boys in general. The velars /k-, k^h-/, labiovelar /k^{wh}-/, and aspirated affricate /ts^h-/ posed more difficulty for boys after age 4;6. The alveolar fricative /s-/ was the last phoneme to be acquired for both boys and

Table 4. Age of acquisition of final consonants and vowels/diphthongs in the present study compared with studies by So and Dodd (1995) and Cheung (1990).

	Final consonants				Vowels/diphthongs	
	Present study	So & Dodd (1995)		Cheung (1990)	Present study	Cheung (1990)[a]
Age group	All	Male	Female	All	All	All
2;6				(-k, -ŋ: by 2;0) -t	i, u, ɛ, ɔ, a, ɪ ɔi, ai, ɐi, ɵy, ou, au	(i, u, ɔ, a, ɐ, y: by 2;0) ɛ (ui, ei, ɔi, ɐi, ɵy, iu, ou, ɐu: by 2;0) au
3;0	-k -n	(-m: by 2;9)	(-p, -k, -m: by 2;9)		y ui, ei, iu, ɛu	œ ai
3;6	-t	(-p, -ŋ: by 3;3)		-m, -p	ʊ, ɐ	
4;0		(-t, -k: by 3;9)	-ŋ	-n	œ ɐu	
4;6	-p, -m, -n, -ŋ		(-n: by 4;3) -t			
5;0	{-k}[b]				ɵ	
Not acquired at 11;7	{-n, -ŋ, -t}[b]					

[a]Cheung (1990) did not differentiate the pairs /i, ɪ/, /u, ʊ/, and /œ, ɵ/. [b]{} represent the age of acquisition when the sound change variants are not considered.

girls at age 6;0. The participants in the present study took half a year longer than those in Cheung's (1990) study and 1 year longer than those in So and Dodd's (1995) study to acquire all the initial consonants.

Final consonants. Table 4 shows that all Cantonese final consonants were acquired by age 4;6 when the substitutions of /-n/ → [-ŋ], /-ŋ/ → [-n], /-t/ → [-k], and /-k/ → [-t] in specific vowel contexts were considered as phonetic variants and the realizations were accepted as correct (see Appendix). If the phonetic variation on final consonants was not taken into account, 90% of the children acquired /-k/ by age 5;0, and /-n, -ŋ, -t/ were not acquired by the oldest participants. The participants in the present study took almost the same amount of time as those in So and Dodd's (1995) study to acquire all final consonants and half a year longer than those in Cheung's (1990) study.

Vowels, diphthongs, and tones. Among the 11 vowels, /i, a, u, ɔ, ɛ, ɪ/ were acquired by age 2;6, whereas /y/ was acquired by age 3;0, /ʊ, ɐ/ by age 3;6, /œ/ by age 4;0, and /ɵ/ by age 5;0 (see Table 4). The only phoneme and allophone pair that were acquired at the same time were /i/ and /ɪ/ (by age 2;6). There was a 1-year lag in the acquisition of the allophones /ʊ/ and /ɵ/, compared with their corresponding phonemes /u/ and /œ/. For diphthongs, /ɔi, ai, ɐi, ɵy, ou, au/ were acquired by age 2;6, /ui, ei, iu, ɛu/ by age 3;0, and /ɐu/ by age 4;0. Tones were acquired early, and the youngest participants (age 2;6) demonstrated 98% of correct tone production. Whereas the children in So and Dodd (1995) acquired all vowels by 2;0 and those in Cheung (1990) acquired all vowels and diphthongs by age 3;0, the children in the present study took 1 to 2 years longer to complete the acquisition. For tones, there is a unanimous agreement that all the nine tones were acquired by the youngest age groups of all the three studies.

Nonadult Realizations and Phonological Patterns of Consonants, Vowels, and Diphthongs

Nonadult realizations and phonological patterns that had 5.0%–9.9% and ≥ 10% occurrence for each phoneme are listed in Figures 3–5. Realizations that occurred less than 5% are not listed.

Initial consonants. Among the initial consonants, /p-, w-, j-, m-, ŋ-/ and syllabic /m/ were acquired early and did not have any nonadult realizations that occurred more than 5%. For the remaining initial consonants, most nonadult realizations occurred until age 3;6 (see Figure 3). The persisting nonadult realizations at age 4;0 were /k^w-/ → [k-], /k^{wh}-/ → [k^h-], /ts-/ → [t-], and /ts^h-/ → [t^h-]; and at age 4;6, the persisting nonadult realization was /s-/ → [ts-]. Dentalization of /s-/ had 5.0%–9.9% occurrence from age 3;6 to 5;6 and persisted between age 6;0 and 11;0 with 1%–2% occurrence.

Figure 4 illustrates the phonological patterns that were frequently found in the children's speech. Common phonological patterns (≥ 10%) were stopping of fricatives and affricates /f-/ → [p-], /s-/ → [t-], /h-/ → [ʔ], /ts-/ → [t-], and /ts^h-/ → [t^h-]; fronting of velars and labiovelars /k-/ → [t-], /k^h-/ → [t^h-], /k^w-/ → [p-], and /k^{wh}-/ → [p^h-]; deaspiration of stops and affricates /p^h-/ → [p-], /t^h-/ → [t-], /k^h-/ → [k-], /k^{wh}-/ → [k^w-], and /ts^h-/ → [ts-]; delabialization of /k^w-/ → [k-] and /k^{wh}-/ → [k^h-]; affrication of /s-/ → [ts-] and /t^h-/ → [ts^h-]; and nasalization of /l/ → [n]. Backing of /t-/ → [k-] and /t^h-/ → [k^h-], deaffrication of /ts^h-/ → [s-], gliding of /l-/ → [j-], and dentalization (lisping) of /s-/ were less common (5.0%–9.9%). Each of these phonological patterns usually appeared from the youngest age group of 2;6 except backing of /t-, t^h-/, affrication of /t^h-/, and deaspiration of /k^{wh}-/, which had 5.0%–9.9% occurrence at age 3;0.

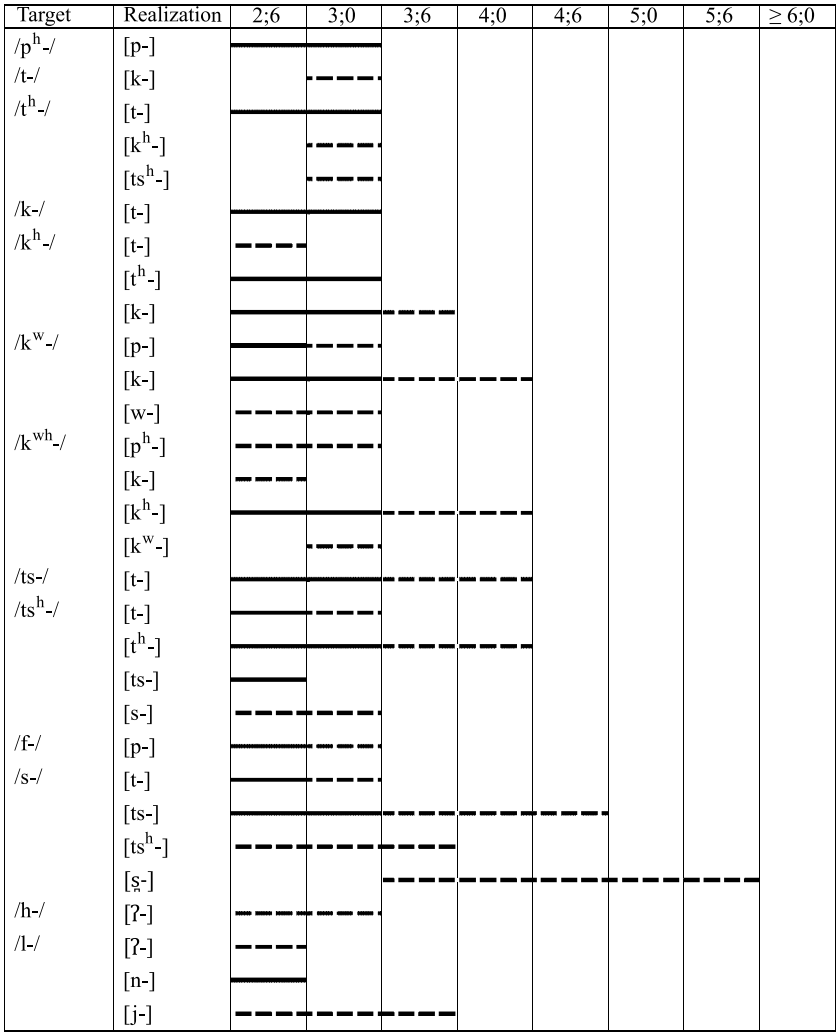
Final consonants. Most nonadult realizations of final consonants resolved between ages 3;0 and 4;0 (see Figure 5). Nonadult realizations of final consonants involved mainly final consonant deletion of the bilabials and nasals /-p, -m, -n, -ŋ/, as well as the backing of the bilabials /-p/ → [-t]/[-k] and /-m/ → [-n].

Vowels and diphthongs. The vowels /i, u, ɛ, ɔ, a, ɪ/ and diphthongs /ɔi, ai, ou, au/ did not have any nonadult realizations that exceeded 5% occurrence (see Figure 5). Common nonadult realizations for the other vowels included spreading of /y/ → [i], lowering of /ɐ/ → [a], and backing of /œ, ʊ, ɵ/ → [ɔ] and /ɵn/ → [ʊŋ]. Most of the nonadult realizations of vowels were resolved by ages 3;6 to 4;6. However, /ɵn/ → [ʊŋ] had a 5.0%–9.9% occurrence that persisted to age 5;6, and continued to have 1%–2% occurrence from age 6;0 to 11;0. The major nonadult realization for diphthongs was the deletion of the second element for /ui, ei, iu, ɛu/. Other realizations were /ɐi/ → [ei], /ɵy/ → [ei], and /ɐu/ → [au]. All nonadult realizations of diphthongs diminished by age 3;0 except /ɐu/ → [au], which had a 5.0%–9.9% occurrence that persisted to age 5;0.

Influence of CV Interactions in Initial Consonant Acquisition

In Table 5, comparisons were made between the occurrence of phonological patterns for initial consonants when combined with front, central, and back vowels. The vowel context was significantly associated with the reduction or increase of the following phonological patterns: (1) Fronting of /k-/ → [t-] and /k^h-/ → [t^h-] was reduced in a back vowel context; (2) stopping of /ts-/ → [t-], /ts^h-/ → [t^h-], and /s-/ → [t-] was reduced in a front vowel context; (3) affrication of /t-/ → [ts-], /t^h-/ → [ts^h-], and /s-/ → [ts-] increased in a front vowel context; and (4) deaffrication of /ts-/ → [s-] and /ts^h-/ →[s-] also increased in a front vowel context. Although different vowel contexts showed no statistically significant difference for the

Figure 3. Summary of nonadult realizations of Cantonese initial consonants. A dashed line indicates that the pattern had 5.0%–9.9% occurrence, and a solid line indicates ≥10% occurrence. Reprinted with the permission from the Child Assessment Service, Department of Health, Hong Kong SAR Government.



occurrence of backing, there was still an observed trend that backing of /t-/ and /t^h-/ was reduced in a front vowel context.

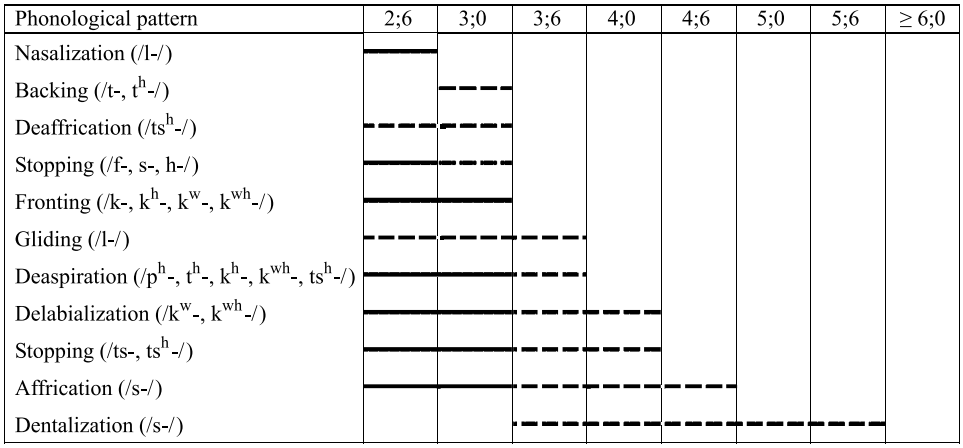
Discussion

This large-scale cross-sectional study of 1,726 Cantonese-speaking children in Hong Kong demonstrated significant age and sex effects for the acquisition of consonants and vowels. Tones were acquired (to 90% correct) by age 2;6, followed by vowels/diphthongs and consonants. There was an interaction effect between vowels and consonants that should be considered during speech sound assessment of Cantonese. We will elaborate on and discuss these findings with respect to previous Cantonese and English studies.

Effect of Age and Sex

Girls acquired initial consonants, final consonants, and vowel/diphthongs significantly earlier than boys. Different from the two previous Cantonese studies by Cheung (1990) and So and Dodd (1995), which found an insignificant trend of sex difference favoring girls in consonant acquisition, the present study obtained a significant sex effect as its large sample size provided sufficient power to achieve this statistical significance. Different from English studies on consonant acquisition—Kenny and Prather (1986) found sex significance favoring girls between ages 3;0 and 5;0; Smit et al. (1990) found this at

Figure 4. Summary of phonological patterns for Cantonese initial consonants. A dashed line indicates that the pattern had 5.0%–9.9% occurrence, and a solid line indicates ≥10% occurrence.



ages 4;0, 4;6, and 6;0; and Dodd et al. (2003) found this at age 5;6–7;0—the present Cantonese study found an Age Group × Sex interaction effect in some younger age groups for initial consonants and final consonants (see Figure 2). The Age × Sex difference in the present study can be explained from the data in Table 3 whereby girls had rapid acquisition of all the initial consonants except /s-/ from ages 2;6 to 4;6. It is expected that the significant sex difference was found during this period of time, whereas boys continued to make progress until age 6;0. As English takes a longer time to acquire than Cantonese, a significant sex difference may appear slightly later. Similar to initial consonants, Cantonese final consonants and vowels/diphthongs also showed a sex effect favoring girls. For

Figure 5. Summary of nonadult realizations of Cantonese final consonants, vowels, and diphthongs. A dashed line indicates that the pattern had 5.0%–9.9% occurrence, and a solid line indicates ≥10% occurrence. Reprinted with the permission from the Child Assessment Service, Department of Health, Hong Kong SAR Government.

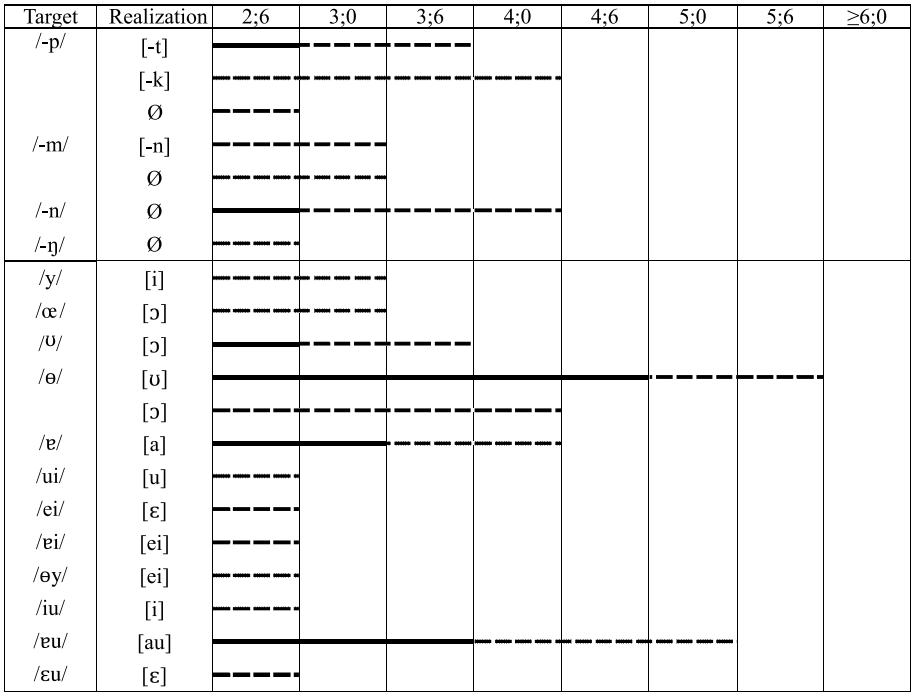


Table 5. Comparison of percentages of nonadult realizations for alveolars and velars in different vowel contexts produced by children from ages 2;6 to 6;0.

Pattern	Nonadult realization	Vowel context			*dfs*	*F*	Partial η^2	*p*
		Front	Central	Back				
Backing	/t-/ → [k-]	[kin]	[kan]	[kou]	2, 84	1.48	.03	.23
		37.2%	53.5%	55.8%				
	/t^h-/ → [k^h-]	[k^him]	[k^hɐi]	[k^hɔŋ]	2, 88	2.54	.06	.08
		40%	53.3%	64.4%				
Fronting	/k-/ → [t-]	[tin]	[tœk]	[tɔn]	2, 170	3.73	.04	.03*
		72.09%	62.8%	52.3%				
	/k^h-/ → [t^h-]	[t^hiu]	[t^hɐm]	[t^hʊk]	2, 212	4.79	.04	.01*
		69.16%	64.49%	49.53%				
Stopping	/ts-/ → [t-]	[tiu]	[tɐi]	[tʊk]	2, 322	38.87	.19	< .01**
		33.3%	72.8%	72.2%				
	/ts^h-/ → [t^h-]	[t^hi]	[t^haŋ]	[t^hɔi]	2, 306	14.44	.09	< .01**
		49.35%	77.92%	59.09%				
	/s-/→[t-]	[ti]	[tam]	[tɔ]	2, 128	13.69	.18	< .01**
		35.4%	75.4%	70.8%				
Deaffrication	/ts-/ → [s-]	[siu]	[sɐi]	[sʊk]	2, 30	3.58	.19	.04*
		81.3%	37.5%	56.3%				
	/ts^h-/ → [s-]	[si]	[saŋ]	[sɔi]	2, 103	3.33	.06	.04*
		65.5%	44.83%	43.1%				
Affrication	/t-/ → [ts-]	[tsin]	[tsan]	[tsou]	2, 63	35.12	.49	< .01**
		81.1%	10.8%	10.8%				
	/t^h-/ → [ts^h-]	[ts^him]	[ts^hɐi]	[ts^hɔŋ]	2, 94	15.0	.24	< .01**
		75.0%	22.92%	35.42%				
	/s-/ → [ts-]	[tsi]	[tsam]	[tsɔ]	2, 270	11.98	.03	< .01**
		53.85%	11.54%	34.62%				

tones, it was reasonable that there was no sex difference, as tones were acquired early with a 98% mastery at age 2;6.

The Acquisition of Cantonese Consonants, Vowels, Diphthongs, and Tones

Children mastered Cantonese in the following order: contrastive tones, vowels/diphthongs, and consonants, which is the same as in the studies by Cheung (1990) and So and Dodd (1995). However, the acquisition of most Cantonese phonemes took slightly longer for the children in the present study than for the children in these two previous studies. There are two possible reasons for the slower rate of acquisition. First, in the present study, consonants were considered to be acquired only when they were correct in all three vowel contexts. This criterion is more stringent than those in the other two studies. So and Dodd (1995) considered a phoneme as acquired if it was produced correctly twice, regardless of the vowel context. In Cheung (1990), there was only one scoring item for each phoneme. Some of the consonants in the two previous studies were not examined in nonfacilitative vowel contexts; for example, there was no /s/+back vowel in either study. As discussed later in this article, CV interaction played a role in the acquisition of consonant production. When nonfacilitative vowel environments are not assessed, children's ability in consonant production may be overestimated, resulting in an earlier acquisition time. Second, children with speech sound disorders were not excluded in the present study, so the sample represented the phonological development of the whole population. It is expected that this inclusive sample would result in a slightly longer acquisition time than when exclusion criteria were applied, such as in Cheung (1990) and So and Dodd (1995).

Initial consonants. The initial consonants /p-, m-, j-/ were acquired by age 3;0 in the present study and were also the earliest to be acquired in Cheung (1990) and So and Dodd (1995). Similar acquisition is also observed in English that /m-, j-/ and the unaspirated voiced /b-/ are among the earliest acquired speech sounds (Dodd et al., 2003; Smit et al., 1990). By age 4;0, the Cantonese /p^h-, f-, h-, w-, ŋ-, t-, l-/ were acquired and were also acquired at a similar or slightly earlier age in the previous two Cantonese studies. In English, similar sounds were acquired between ages 3;0 and 3;6 (Dodd et al., 2003; Smit et al., 1990), except that /l-/ was acquired at age 5;0 by the girls and at age 6;0 by the boys in Smit et al. (1990). If the

English /l/ in the other word positions are considered, the sound was acquired even later, from age 5;6 to 8;0 (Smit et al., 1990). The heavy functional load of the Cantonese /l-/ may explain its early acquisition. The phoneme occupied the third rank in the frequency of occurrence among Cantonese initial consonants, as it has almost replaced /n-/ in the process of sound change (Ng & Kwok, 2004). Such a frequent exposure may have facilitated the acquisition of the phoneme.

In the present study, /t^h-, ts-, k-, k^h-, k^w-, and k^{wh}-/ were acquired by age 5;0. Participants in So and Dodd (1995) took from age 2;6 to age 5;0 to acquire these alveolars (including the affricate /ts/) and velars, whereas participants in Cheung (1990) took a shorter time, from age 3;0 to age 4;0. In English, velars /k-, ɡ-/ were acquired early, by age 3;6 (Dodd et al., 2003; Smit et al., 1990). A possible reason why it took longer for participants to acquire Cantonese velars than English velars may be because /k^h-, k^w-, k^{wh}-/ have a low frequency of occurrence in Cantonese. Although /k-/ occupied the first position of the initial consonant frequency ranking in Ng and Kwok (2004), /k^h-, k^w-, k^{wh}-/ had positions of 14th, 17th, and 19th, respectively, in the ranking. As a result, the low functional load of most velars may play a part in the slower rate of acquisition.

The last phonemes to be acquired in the present study were /ts^h-/ and /s-/ by the age of 6;0. Cheung (1990) and So and Dodd (1995) also reported that these two sounds are among the most difficult. In English, /s/ and /ʃ-/ are reported to be among the latest developing sounds in a number of English studies (see, e.g., McLeod, 2013, for a review). Such a consistent pattern indicates that the inherent physiological requirements in producing these phonemes are complicated. For example, for the production of /s/, the generation of the central groove and the maintenance of a continuous airflow require fine-tuned articulatory gestures and good coordination.

Despite the slight variations in acquisition time, the present study followed the same initial consonant acquisition order as did the two previous Cantonese studies and other English studies; that is, the place order of bilabials followed by alveolars and then velars and the manner order of nasals, plosives, and approximants followed by fricatives and affricates. From these data, it appears that both the articulatory ease and the functional load of a phoneme in a language may contribute to its acquisition rate.

Final consonants. For final consonants, there was an interesting phenomenon of reverse acquisition rate compared with their initial counterparts; that is, final /-k/ (age 3;0) and /-t/ (age 3;6) were acquired before initial /k-/ (ages 4;0–5;0) and /t-/ (ages 3;6–4;0), whereas final /-p, -m, -ŋ/ (age 4;6) were acquired much later than initial /p-/ (age 2;6), /m-/ (age 3;0), and /ŋ-/ (ages 2;6–3;6). Initial /n-/ was realized as [l-] (an acceptable phonetic variant; see Appendix), so it cannot be compared with /-n/. This reverse pattern was also found in Cheung (1990), except for /-ŋ/. In So and Dodd (1995), the boys and the girls had different acquisition patterns of final consonants, except that the final bilabials also showed later acquisition than their initial counterparts for both sexes. Initial bilabials are acquired early, as lip closure for the sounds coincides with the position when the lips are at rest, whereas final bilabials require lip closure of appropriate timing and take more time to accomplish. In addition, /-p/ and /-m/ have a relatively low frequency of occurrence in Cantonese, making up only 2.54% and 12.51%, respectively, of the total final consonant occurrence (last and third-to-last in ranking; Ng & Kwok, 2004). Thus, it seems that articulatory effort and the functional load of a phoneme again may contribute to its acquisition rate. English final consonants do not show a clear pattern of acquisition (Smit et al., 1990) and cannot be compared with Cantonese.

The present data accepted the phonetic variants of /-n, -ŋ, -t, -k/ in specific vowel contexts (see Appendix). The data demonstrate that this was an appropriate consideration; otherwise, the acquisition of /-n, -ŋ, -t/ would have gone beyond the oldest age group of 11 years old (see Table 3). The variation in final consonants is vowel dependent (To et al., 2012), and most of the target items in the Cheung study were attached to vowels that did not facilitate the production of the variants; for example, /-ŋ/ and /-k/ were attached to a high-back vowel in /fʊŋ/ ("wind") and /ʊk/ ("house"), and /-t/ was attached to a high-front vowel in /syt/ ("ice"). Phonetic variants for final consonants were not recorded in Cheung (1990) or in So and Dodd (1995).

Vowels. Cantonese vowels rarely have been studied in depth because they are acquired early. So and Dodd (1995) reported that all vowels were acquired by age 2;0, whereas Cheung (1990) reported that they were acquired from age 2;0 to 3;0. In the present study, the vowel acquisition spanned from age 2;6 to 5;0. In the present study, the six vowels /i, ɪ, u, ɔ, a, ɛ/ were acquired early, as they were in Cheung (1990) and Stokes and Wong (2002), in which the ranked order of vowel accuracy at 24 to 27 months was /ɛ/ > /ɔ/ > /a/ = /i/ > /ɐ/ > /œ/ > /u/ > /y/. Among these six vowels, /i, a, u, ɔ/ were also found to be acquired early in English-speaking children and in children with speech sound disorders (Stoel-Gammon & Herrington, 1990; Pollock & Keiser, 1990). Donegan (2002) suggested that the optimal palatal (front) vowel is [i], the optimal labial (round) vowel is [u], and the optimal sonorant vowel is the nonpalatal and nonlabial [a], and that these three vowels are acquired early and are favored among the world's languages. The vowels that were acquired subsequently in the present study were /y/ (age 3;0), /ɐ, ʊ/ (age 3;6), /œ/ (age 4;0), and /ə/ (age 5;0). Cheung (1990) and Stokes and Wong (2002) did not study allophones separately. Cheung's (1990) vowels were

acquired early because the allophones /ɪ, ʊ, ɵ/ were not examined. Stokes and Wong (2002) included allophones /ɪ, ʊ, ɵ/ into the phonemes /i, u, œ/, respectively, and both /œ/ and /u/ were acquired late. Additional reasons for the delay in acquisition of /y/ may include the presence of the lip-rounding element in /y/. The short vowels /ɐ, ɪ, ʊ, ɵ/ must be followed by a final consonant. Any problem with the final consonant—for example, deletion or substitution—may cause instability in the preceding vowel. Last of all /ɐ, ʊ, œ, ɵ/ do not occupy any of the optimal palatal, optimal labial, or optimal sonorant positions as suggested by Donegan (2002), and this may lead to a later acquisition. The allophone /ɵ/ possesses all three characteristics—lip rounding, nonoptimal position, and occurring in closed syllables—and this may explain why it was acquired last.

Diphthongs. Children in Cheung (1990) acquired all diphthongs between ages 2;0 and 3;0, and children in the present study acquired 10 of 11 diphthongs by age 3;0. By age 4;0, the last diphthong /ɐu/ was acquired. Although /ɐ/ is not an early acquired vowel, /ɐi/ was acquired early, at age 2;6, in the present study and did not share the same difficulty as /ɐu/. The test item for /ɐu/ in the present study was /ŋɐu/ ("cow"). As 78% of the children used the zero initial instead of the standard form /ŋ-/, /ɐu/ became a syllable not bound by any initial or final consonant. It is speculated that deletion of the initial consonant might have increased the syllable duration of the first element in /ɐu/, so the realization was perceived as [au]. Whether this /ɐu/-to-[au] error was item specific or a general problem applied to other /-ɐu/ words requires further investigation.

Tones. In the present study, participants had acquired Cantonese tones by age 2;6, and this early acquisition is also supported by Cheung (1990) and So and Dodd (1995).

Phonological Patterns

Initial consonants. Among the six common phonological patterns (≥ 10% occurrence) of stopping, fronting, deaspiration, delabialization, affrication of /s-/, and nasalization of /l-/ in initial consonants, all were reported in Cheung (1990), and all except nasalization were reported in So and Dodd (1995). In both the Cheung (1990) and the So and Dodd (1995) studies, all four less common patterns (5.0%–9.9% occurrence) of backing, deaffrication of /ts^h-/, gliding, and dentalization were found in a small number of children. These three Cantonese studies agree regarding the presence of the phonological patterns, but it is difficult to make a direct comparison of the occurrence rate because Cheung (1990) showed the number of occurrences, whereas So and Dodd (1995) showed the percentage of children making errors.

Stopping of the fricatives /f-, s-, h-/ and affricates /ts-, ts^h-/ to their homorganic plosives diminished remarkably by ages 3;6 and 4;6, respectively. Stopping is also common for young, English-speaking children (see, e.g., Dodd et al., 2003; Grunwell, 1987; Smit, 1993). For example, in Smit (1993), substitution of /f-, s-/ with [p-, b-, t-] made up 5%–15% of occurrence and substitution of /ʃ-/ with [t-, d-] showed 15%–30% of occurrence in the ages 2;0–3;0. The pattern became "occasional" or "rare" in the years after.

Fronting in the present study occurred in the velar series /k-, k^h-, k^w-, k^{wh}-/ and was mainly found in the 2;6 and 3;0 age groups. Fronting in Cantonese disappears slightly earlier than that in English. For example, substitution of /k-, ɡ-/ by [t-, d-] in Smit (1993) had 5%–15% occurrence in the ages 2;0–3;6.

Deaspiration in Cantonese involves the inability to produce the aspiration in the voiceless plosives and affricate, (e.g., /p^h-/ → [p-]). Its only difference from the English process of voicing—for example, /p-/ → [b-]—is that no voicing element is involved. In the present study, deaspiration was commonly found in /ts^h-/ in the 2;6 group and in /p^h-, t^h-, k^{wh}-/ in the 2;6–3;0 groups, and it extended to age 3;6 in /k^h-/. Smit (1993, p. 536) also used the term *deaspiration* instead of voicing in her English study, whereby /p-/ → [b-], /t-/ → [d-], and /k-/ → [ɡ-] were found occasionally (< 4%) until age 5;0 and became rare (< 3%) in the ages after that, whereas /ʃ-/ → [ʤ-] had a rare occurrence from age 3;6 to age 9;0. It is common in both English and Cantonese that aspiration requires time to acquire.

Delabilization of /k^w-/ → [k-] and /k^{wh}-/ → [k^h-] is specific to Cantonese, as English does not have labiovelars. Problem in producing the lip-rounding element [^{w}] was common in participants ages 2;6–4;0, making the two labiovelars among the last acquired phonemes in Cantonese.

Affrication of /s-/ → [ts-] → [ts^h-] was found across the age groups of 2;6–4;6 in the present study, but this process has not been recorded in English (Dodd et al., 2003; Grunwell, 1987; Ingram, 1976; Smit, 1993). In the present study, affrication was more common than the stopping of /s-/ to [t-], which mainly appeared in participants ages 2;6–3;0. In the clinical observation of the Cantonese-speaking authors of the present study, [ts-] appears to be a more mature realization than [t-] in the acquisition of /s/, as many children produce the following transition of [t-] → [ts-] → [s-]. The Cantonese phonemes /s-, ts-, ts^h-/ share the same place of articulation. It is speculated that when a child tries to produce the fricative /s-/, adding a fricative element (i.e., [+continuant]) to [t-] will result in a homorganic affricative [ts-] or [ts^h-]. Further effort will be needed to remove the [-continuant] feature in [ts-] and [ts^h-] to obtain the adultlike /s-/. Affrication of /s-/ in English may be rare because English does not have an affricate that is homorganic to /s/.

Nasalization of /l-/ to [n-] frequently occurred only in the youngest group (age 2;6) and diminished in the older groups. The Cantonese /l-/, which is an apical dentalveolar

(Zee, 1999), is a difficult sound for young children, as it requires only the tongue tip touching the alveolar region. It is easy for young children to realize /l-/ as [n-], as /n/ is an apicolaminal dentialveolar (Zee, 1999), which allows a larger anchorage of the tongue in both the apical and laminal area on the hard palate and upper teeth, and /n/ is also a frequently occurring final consonant in Cantonese (Ng & Kwok, 2004). Smit (1993) also reported that the same phonological pattern was found to be "occasional" (< 4%) or "rare" (< 3%) among the participants from age 2;0 to 9;0.

Phonological patterns at 5.0%–9.9% occurrence in the present study included backing, deaffrication, gliding, and dentalization. Backing of /t-/ to [k-] and /t^h-/ to [k^h-] in Cantonese occurred mainly in the 3;0 group and diminished sharply in the older age groups. In English, however, it is considered as an atypical or idiosyncratic error (Dodd et al., 2003; Grunwell, 1987; Ingram, 1976). Smit (1993) reported a < 3% occurrence of [k-] for /t-/ in ages 2;0–3;6 and [ɡ-] for /d-/ in ages 3;6–9;0. The frequent occurrence of /k-/ and /-k/ (ranked first and third, respectively, in Ng & Kwok, 2004) in Cantonese may reinforce the realization of the velars as substitutions for alveolar plosives.

Deaffrication of /ts^h-/ to [s-] was found mainly in children ages 2;6–3;0. Although So and Dodd (1995) found it a common process in both /ts-/ and /ts^h-/, /ts-/ → [s-] was not frequently found in the present study, and it was not recorded in Cheung (1990). Smit (1993) had a finding comparable to that in the present study that a 5%–15% occurrence of English /ʧ-/ → [ʃ-] was found in children between ages 2;0 and 4;6, but /ʤ-/ → [ʒ-] was only occasionally (< 4%) or rarely (< 3%) found in all participants.

Gliding of /l-/ in the present study had 5.0%–9.9% occurrence in young children ages 2;6–3;6 and was more extensively used than nasalization in /l-/. In Smit (1993), English gliding of /l-/ had an occurrence of < 4% in children ages 2;0–5;6.

Dentalization of /s-/ rose to a 5.0%–9.9% occurrence in children ages 3;6–5;6 and persisted (1%–2% occurrence) from age 6;0 to age 11;0. Dentalization was regarded as a nonadult realization of Cantonese initial consonants that persisted through primary school. In English, Smit (1993) reported an even greater occurrence of dentalization in /s/. For the English /s-/, there was a 15%–30% occurrence between the ages of 2;0 and 5;6, and the occurrence dropped to 5%–15% between the ages 6;0 and 9;0. Smit (1993) found that, by age 6;0, virtually all error types in English /s, z/ were dental. In both languages, the relatively constant percentage of occurrence of dentalization in the older age groups suggests a long-term persistence, even to adulthood. As the dental [s-] is close to the target /s-/ in Cantonese and does not arouse any phonemic confusion (in English, it can be confused with /θ/), Cantonese parents and children may not be aware that the dental [s] is a problem production. Therefore, it is recommended that this phonological pattern be managed clinically once it is identified.

Final consonants. The six final consonants in Cantonese were realized differently from their initial counterparts. The main pattern was final consonant deletion, which was mostly found in /-p, -m, -n, -ŋ/ among children ages 2;6–4;0. The process rarely occurred in Cantonese /-k, -t/; this is different from English studies in that final consonant deletion is common to all English final consonants (Smit, 1993), making these two final consonants the earliest to acquire. It is uncertain why /-k, -t/ are relatively free from deletion.

Vowels and diphthongs. The vowels /i, u, ɛ, ɔ, a, ɪ/ and diphthongs /ɔi, ai, ou, au/ were acquired early. Phonological patterns were found in children's nonadult realizations of the vowels /y, œ, ʊ, ɵ, ɐ/ and diphthongs /ui, ei, ɐi, ɵy, iu, ɐu, ɛu/. For vowels, delabialization, backing, and lowering were the three major patterns. Delabialization of /y/ to [i] had a 5.0%–9.9% occurrence in children ages 2;6–3;0. It was also the major nonadult realization for /y/ in Stokes and Wong (2002), who claimed that lip rounding requires motoric competence and leads to the acquisition complexity of the vowel. In addition to that, substituting /y/ with [i] increases the sonority of the sound, and delabialization is common in typical vowel acquisition in English-speaking children (Donegan, 2002). Backing also is a common vowel error in English-speaking children (Pollock, 1994). In Cantonese, /œ/ was frequently backed to [ɔ] in children ages 2;6–3;0, whereas /ɵ/ → [ɔ] extended to age 4;0. Lowering occurred in the short vowel /ʊ/, which was realized as [ɐ] in children under age 4;0, and in the vowel /ɐ/, which was realized as [a] in children under age 4;6. Donegan (2002) suggested that lowering increased sonority and preferentially affected lax vowels, which precisely describes the behavior of Cantonese /ʊ/ and /ɐ/. Comparison of phonological patterns in Cantonese and English vowels demonstrated a universal trend in typical vowel acquisition. Allophones of vowels were acquired at different rates and demonstrated different phonological patterns; it is a correct consideration to examine vowels and their allophones separately to get a full view of vowel acquisition.

Simplification of diphthongs was the major phonological pattern found in /ui, ei, iu, ɛu/ in which the second element of the diphthong was deleted. For realizations that were maintained as diphthongs—namely, /ɐi/→[ei], /ɵy/→[ei], and /ɐu/→[au]—the patterns involved lengthening the production time of the first short vowel to become a long vowel and losing the lip-rounding feature in /ɵy/. It is interesting to note that diphthongs were acquired relatively earlier than were pure vowels and that most of the phonological patterns in diphthongs had < 10% occurrence in the youngest age group. There

is limited information regarding the development of English diphthongs for comparison.

Influence of CV Interactions in Initial Consonant Acquisition

Vowel contexts played a role in the realization of the place and manner of a consonant in the phonological acquisition of Cantonese. A back vowel significantly reduced the chance of a velar (back) plosive being realized as a front consonant. A high-front vowel was also observed to reduce the chance of an alveolar plosive being realized as a back consonant, although the results were not statistically significant. These two phonological patterns involved a change in the place of articulation and were consistent with the frames then content theory (Davis & MacNeilage, 1995; Davis et al., 2002) that velars with a more posterior tongue position can best accommodate the back vowels and that coronals with a more anterior tongue position can better accommodate high-front vowels.

Vowel contexts also played a role in the manner of consonant realization. In the stopping process where the affricates /ts, ts^h-/ and fricative /s-/ were realized as alveolar plosives [t-], [t^h-] and [t-], respectively, a high-front vowel reduced plosive realizations more significantly than did central and back vowels. If the patterns were analyzed with reference to the articulation position of the consonants and vowels, a proximity facilitation effect can be shown. Cantonese /t-, t^h-/ are apicolaminal dentialveolars and are produced with a more anterior part of the tongue and palate than are the laminal-alveolars /ts-, ts^h-, s-/ (Zee, 1999). The tongue palate involvement of the high-front vowel /i/ in Cantonese is, in fact, closer to /ts-, ts^h-, s-/ than to /t-, t^h-/. Thus, a high-front vowel context significantly reduced the realization of [t-, t^h-, t-] in the stopping of /ts-, ts^h, s-/, respectively. On the other hand, as the combination of /ts-, ts^h-, s-/ with a high-front vowel is favored, a front vowel context increased the ease of the affrication of /t-, t^h-, s-/ to [ts-, ts^h, ts-], respectively, as well as the deaffrication of /ts-, ts^h-/ to [s-]. This evidence suggests that vowel contexts not only reduce but also facilitate the realization of certain nonadult productions in phonological development.

CV interaction has been found to be one of the many factors that affect the acquisition of Cantonese consonants in the present study as well as in children with speech sound disorders (Cheung & Abberton, 2000). Therefore, it is recommended that a Cantonese speech assessment should include consonant stimuli combined with front, central, and back vowels to assess the stability of consonant acquisition. In addition, examining vowels and their allophones separately will refine the differentiation of CV interactions. For example, high-front vowels such as /i/ may contribute to an interaction, whereas /ɪ/ may not. When assessing children with speech sound disorders, alveolars with high-front vowels or velars with back vowels can be used to examine stimulability. However, as only three vowels in front, central, and back positions for each consonant were examined in the present study, an exhaustive investigation of consonant acquisition in all vowel contexts in both typical and disordered speech development may help reveal the interaction in full detail.

Conclusion

This population study of the speech of Hong Kong Cantonese-speaking children (N = 1,726) presents a full picture of phonological acquisition and related error patterns of Hong Kong Cantonese. By accepting ongoing sound change variants, studying vowels and their allophones separately, and examining the effect of CV interaction, a precise description of the acquisition of Hong Kong Cantonese has been obtained that can subsequently assist with the identification and management of children with speech sound disorders. The present study shows that Cantonese acquisition shares some similarities with English acquisition and has its specific characteristics as well. This large-scale research on Hong Kong Cantonese may serve as a basis for studies of child speech in other Chinese dialects and dialect varieties.

Acknowledgments

The third author acknowledges support from Australian Research Council Future Fellowship (FT0990588). We gratefully acknowledge the Child Assessment Service of the Government of Hong Kong SAR, the City University of Hong Kong, and all the SLPs who participated in this project. We thank the schools, the children, and their parents who helped us with this research.

References

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. Berlin, Germany: Mouton de Gruyter.

Census and Statistics Department. (2007). Proportion of population aged 5 and over able to speak selected languages/dialects 1996, 2001 and 2006. In *2006 Population by-census*. Retrieved from www.bycensus2006.gov.hk/FileManager/EN/Content_981/all4e.xls.

Cheung, P., & Abberton, E. (2000). Patterns of phonological disability in Cantonese-speaking children in Hong Kong. *International Journal of Language and Communication Disorders, 35,* 451–473.

Cheung, P. S. P. (1990). *The acquisition of Cantonese phonology in Hong Kong: A cross-sectional study.* Unpublished undergraduate thesis, University College London.

Cheung, P. S. P., Ng, A., & To, C. K. S. (2006). *Hong Kong Cantonese Articulation Test*. Hong Kong SAR, People's Republic of China: Language Information Sciences Research Centre, City University of Hong Kong.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 38,* 1199–1211.

Davis, B. L., MacNeilage, P. F., & Matyear, C. (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production. *Phonetica, 59,* 75–107.

Dodd, B., Holm, A., Zhu, H., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics, 17,* 617–643.

Donegan, P. (2002). Normal vowel development. In M. J. Ball & F. E. Gibbon (Eds.), *Vowel disorders* (pp. 1–36). Woburn, MA: Butterworth-Heinemann.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers, 28,* 1–11.

Grunwell, P. (1987). *Clinical phonology*. Baltimore, MD: Williams & Wilkins.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53–69.

Ingram, D. (1976). *Phonological disability in children*. London, England: Edward Arnold.

Kenny, K. W., & Prather, E. M. (1986). Articulation development in preschool children: Consistency of productions. *Journal of Speech and Hearing Research, 29,* 29–36.

Lewis, M. P. (2009). *Ethnologue: Languages of the world* (16th ed.). Dallas, TX: SIL International. Retrieved from www.ethnologue.com.

McFadden, T. U. (1996). Creating language impairments in typically achieving children: The pitfalls of "normal" normative sampling. *Language, Speech, and Hearing Services in Schools, 27,* 3–9.

McLeod, S. (2013). Speech sound acquisition. In J. E. Bernthal, N. W. Bankson, & P. Flipsen, Jr. (Eds.), *Articulation and phonological disorders: Speech sound disorders in children* (7th ed., pp. 58–113). Boston, MA: Allyn & Bacon.

Ng, M. L., & Kwok, I. C. L. (2004). Frequency of occurrence of Cantonese word initials, finals and tones. *Asia Pacific Journal of Speech, Language, and Hearing, 9,* 189–199.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology, 15,* 247–254.

Pollock, K. E. (1994). Assessment and remediation of vowel misarticulations. *Clinics in Communication Disorders, 4,* 23–37.

Pollock, K. E., & Keiser, N. J. (1990). An examination of vowel errors in phonologically disordered children. *Clinical Linguistics and Phonetics, 4,* 161–178.

Smit, A. B. (1993). Phonological error distributions in the Iowa–Nebraska Articulation Norms Project: Consonant singletons. *Journal of Speech and Language Research, 36,* 533–547.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Language Disorders, 55,* 779–798.

So, L. K. H. (1993). *Cantonese Segmental Phonology Test*. Hong Kong SAR, People's Republic of China: Bradford.

So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language, 22,* 473–495.

Stoel-Gammon, C., & Herrington, P. B. (1990). Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics and Phonetics, 4,* 145–160.

Stokes, S. F., & To, C. K. S. (2002). Feature development in Cantonese. *Clinical Linguistics and Phonetics, 16,* 443–459.

Stokes, S. F., & Wong, I. M. (2002). Vowel and diphthong development in Cantonese-speaking children. *Clinical Linguistics and Phonetics, 16,* 597–617.

Sussman, H. M., Duder, C., Dalston, E., & Cacciatore, A. (1999). An acoustic analysis of the development of CV coarticulation: A case study. *Journal of Speech, Language, and Hearing Research, 42,* 1080–1096.

To, C. K. S., Cheung, P. S. P., & McLeod, S. (2012). *Phonetic variations and sound changes in Hong Kong Cantonese and implications for speech sound assessment*. Manuscript submitted for publication.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565–604). Timonium, MD: York Press.

Wong, W.-Y., W., & Stokes, S. F. (2001). Cantonese consonantal development: Towards a nonlinear account. *Journal of Child Language, 28,* 195–212.

Zee, E. (1999). Chinese (Hong Kong Cantonese). In *Handbook of the International Phonetic Association* (pp. 58–60). Cambridge, England: Cambridge University Press.

Appendix. The nine phonetic variations (sound changes) that were accepted as correct pronunciations in the present study (based on To, Cheung, & McLeod, 2012).

Phonetic Variants of Ongoing Sound Changes Related to Initial Consonants.

1. /n-/ → [l-]: Initial alveolar nasal /n/ is produced as the approximant [l].
2. /ŋ-/→Ø-: Initial velar nasal /ŋ-/ is omitted.
3. Ø- → [ŋ-]: Zero initial Ø- is produced as initial velar nasal /ŋ/.
4. /k^wɔ-/ → [kɔ-]: Initial labialized velar plosive /k^w/ is produced as the velar plosive [k] in the context of the /ɔ/ vowel.
5. Syllabic /ŋ/ → [m]: Initial velar nasal /ŋ/ is produced as [m] in syllabic consonants (i.e., when the consonant forms the entire syllable).

Phonetic Variants of Ongoing Sound Changes Related to Final Consonants.

6. /-ŋ/ → [-n]: Final velar nasal /-ŋ/ is produced as alveolar nasal /-n/ in all vowel contexts except high vowels including /ɪ/ and /ʊ/.
7. /-n/ → [-ŋ]: Alveolar nasal /-n/ is produced as the final velar nasal /-ŋ/ in the vowel contexts of /ɔ/ and /a/.
8. /-k/ → [-t]: /k/ is produced as a final alveolar plosive /-t/ in the vowel contexts of /ɔ/ and /œ/.
9. /-t/ → [-k]: /t/ is produced as a final velar plosive in vowel contexts of /ɐ/ and /ɔ/.

Note. The rime /ʊŋ/ was examined in a child group (age 11 years), but not the adult group, and the occurrence of nonstandard production was 0%.