Suprasegmental features are not acquired early: Perception and production of monosyllabic Cantonese lexical tones in four- to six-year-old pre-school children

Puisan Wong[1] & Carrie Tsz-Tin Leung[2]

[1]Corresponding author. Division of Speech and Hearing Sciences, Faculty of Education, The University of Hong Kong, Pokfulam, Hong Kong. Email: pswResearch@gmail.com

[2]Division of Speech and Hearing Sciences, Faculty of Education, The University of Hong Kong, Pokfulam, Hong Kong. Email: eirra328@connect.hku.hk

Acknowledgements

This research was supported by funding from the Seed Funding Programme for Basic Research from the University of Hong Kong to the first author.

Abstract

Purpose: Previous studies reported that children acquire Cantonese tones before three years of age, supporting the assumption in models of phonological development that suprasegmental features are acquired rapidly and early in children. Yet, recent research found a large disparity in the age of Cantonese tone acquisition. This study investigated Cantonese tone development in four- to six-year-old children.

Method: Forty-eight four- to six-year-old Cantonese-speaking children and 28 mothers of the children labeled 30 pictures representing familiar words in the six tones in a picture naming task and identified pictures representing words in different Cantonese tones in a picture pointing task. To control for lexical biases in tone assessment, tone productions were low-pass filtered to eliminate lexical information. Five judges categorized the tones in filtered stimuli. Tone production accuracy, tone perception accuracy and correlation between tone production and perception accuracy were examined.

Results: Children did not start to produce adult-like tones until five and six years of age. Four-year-olds produced none of the tones with adult-like accuracy. Five- and six-year-olds attained adult-like productions in two (T5 and T6) to three (T4, T5 and T6) tones, respectively. Children made better progress in tone perception and achieved higher accuracy in perception than in production. However, children in all age groups perceived none of the tones as accurately as adults, except that T1 was perceived with adult-like accuracy by six-year-olds. Only weak association was found between children's tone perception and production accuracy.

Conclusions: Contradicting to the long-held assumption that children acquire lexical tone rapidly and early before the mastery of segmentals, this study found that four- to six-year-old children have not mastered the perception or production of the full set of Cantonese tones in familiar monosyllabic words. Larger development was found in children's tone perception

than tone production. The higher tone perception accuracy but weak correlation between tone perception and production abilities in children suggested that tone perception accuracy is not sufficient for children's tone production accuracy. The findings have clinical and theoretical implications.

Introduction

Cantonese, is a tone language spoken by 7.3 million people world-wide (Simons & Charles, 2017). Like other tone languages, it uses lexical tone, the change of pitch at the syllable level, to distinguish word meanings (Yip, 2002). Various populations have been found to have special difficulties with tone perception and production, such as children with cochlear implants (Ciocca, Francis, Aisha, & Wong, 2002), cerebral palsy (Chen, Ni, Kuo, & Hsu, 2012) and dyslexia (Cheung et al., 2009). Despite the importance of lexical tones, due to large discrepancies in previous studies, no conclusive results are available on the development of tones in Cantonese-speaking children, compromising a clear understanding of typical and atypical speech development in children, proper evaluation and treatment of children with tone perception and production difficulties, and the establishment of universal theories of phonological acquisition.

Cantonese Tones

Among tone languages, Cantonese has one of the most complex tone systems. It consists of six non-checked/ non-entering tones and three checked / entering tones (Wong & Chan, 2018). The non-checked tones include three level tones (i.e. High-Level (HL, T1), Mid-Level (ML, T3), and Low-Level (LL, T6)), two rising tones (i.e. High-Rising (HR, T2) and Low-Rising (LR, T5)) and a falling tone (Low Falling (LF, T4)). The three level tones differ from one another in the height of the fundamental frequency (F0) while the two rising tones differ from each other in both F0 height and F0 slope. Thus, to master accurate perception and production of tones, children need to be able to perceive and produce both the F0 heights and shapes of different tones accurately. Cantonese also has three checked tones that occur in closed syllables with a final plosive consonant. Modern Cantonese phonology considers the checked tones allotones of the three non-checked level tones (Mok, Zuo, & Wong, 2013; So & Dodd, 1995). Previous developmental studies did not examine the checked tones. A recent

detailed acoustic study showed that the duration and the F0 contours of the checked tones are different from the three level non-checked tones (Wong & Chan, 2018).Thus, the checked tones were excluded from this study. Figure 1 shows the F0 contours of the six non-checked tones.

Figure 1

Age of Acquisition of Lexical Tones in Children

Most previous studies on children's production of lexical tones reported that children acquire the six tones at around 2 to 3 years of age (e.g., Tuaycharoen (1977) on Thai tones; Hua & Dodd (2000) on Mandarin tones; and To, Cheung, & McLeod (2013) on Cantonese tones). The findings support the tenet in models of phonological development that children master suprasegmental units much earlier than segmental units (Dehaene-Lambertz & Houston, 1998; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984).

Cantonese Tone Production in Children

Despite the complex tone system in Cantonese, most studies on children's production of Cantonese tones reported that children mastered the production of the six tones before three years of age. Three of these studies were case studies with a small sample size. Tse (1978) conducted a longitudinal study on a Cantonese-speaking child aged 2;8 and reported that the child mastered Cantonese tones at the onset of the study, suggesting that children acquire the tonal system before 3;0 (J. K.-P. Tse, 1978). So & Dodd (1995) carried out a longitudinal study on four younger children aged 1;2 - 2;0 and reported that children acquired all the tones by 2;0. Tse (1992) tracked the tone development of a single child from 1;2 - 3;0 and supported that the tone system was acquired before the consonant system, with signs of nearing completion of all tones at 3;0 (A. C.-Y. Tse, 1992).

Findings of more recent large-scale studies also supported early acquisition of Cantonese tone production. So and Dodd (1995) examined tone and segmental productions in 268 two- to six-year-old children. Only two tone errors were found in one four-year-old and three tone errors were found in one five-year-old. The authors concluded that the children acquired the tones and vowels at two years of age. Using a similar method, To et al. (2013) studied tone and segmental productions in 1726 children aged 2;4 - 12;4 in Hong Kong. Eleven speech therapists rated children's accuracy of tones during the administration of the Hong Kong Cantonese Articulation Test (HKCAT). The results showed that children in the youngest age group produced the tones with 98% accuracy and the authors concluded that all children produced all tones correctly by 2;6.

Lexical expectation of judges and the presence of environmental and contextual cues during tone judgment might have interfered judges' tone assessment, leading to the findings of early acquisition of Cantonese tone production in the aforementioned studies. Judges in these studies rated the tones in unfiltered speech and were not blinded to the target tone or the lexical item. Many studies asked the children to imitate the productions of the experimenters when the child did not know the lexical items (e.g., So & Dodd, 1995; Tse, 1978; Hua & Dodd, 2000; To et al., 2013). Some studies rated children's tones in spontaneous speech with the presence of the environmental, syntactic and semantic cue (So & Dodd, 1995; A. C.-Y. Tse, 1992). Prior research has shown that when lexicality expectation and context of word stimuli are present, listeners may not be able to detect differences in speech stimuli (Davis & Johnsrude, 2007; Oller & Eilers, 1975). Moreover, none of these studies performed acoustic analysis to evaluate the consistency between the physical properties of the produced tones and the judges' tone categorization in unfiltered speech. Furthermore, none of the studies, except To et al., (2013), examined inter- and intra-judge reliability.

Studies on children's Mandarin tone acquisition demonstrated large effect of lexical expectation on tonal judgment. Studies in which judges rated children's tones in unfiltered speech with lexical information available reported that children acquired the four Mandarin tones by 2;0 (e.g., (Clumeck, 1980; Hua, 2002; Hua & Dodd, 2000). On the other hand, studies that asked judges to categorize children's and adults' tones in filtered speech found that even 5-year-old children did not produce the four Mandarin tones in monosyllabic words with adult-like accuracy (Wong, 2013).

Two studies on children's production of Cantonese tones that controlled lexical bias in tone rating reported later acquisition of Cantonese tones. Barry and Blamey (2004) investigated tone production of eight 4 to 6-year-old normal children, sixteen 4 to 11-year-old cochlear implant users and five normal adults. A non-native Cantonese speaker transcribed the tones based on the pitch information, which might have reduced lexical biases of the judge. In addition, onset and offset frequencies of the tones were measured and compared among the three groups of speakers. The results revealed that though adults' tone productions were not error-free, tone accuracy of typically-developing children was not adult-like. Even 6-year-old children were still learning to normalize for pitch level differences in tone productions and demonstrated confusions among tones with similar shapes of pitch contours, such as the three level tones and the two rising tones (Barry & Blamey, 2004).

Wong, Fu, and Cheung (2017) examined monosyllabic Cantonese tone production of three-year-old children using the methods in the Mandarin studies that controlled lexical expectation (Wong, 2012b; Wong, Schwartz, & Jenkins, 2005) and performed acoustic analysis to compare the acoustic properties of adults' and children's Cantonese tones that were correctly and incorrectly perceived by the judges. Tone productions were collected from 19 children and their mothers. Five native Cantonese speakers who were blinded to the stimuli were recruited as judges to identify the tones in filtered stimuli. Overall, adults' tones

were identified at around 98% accuracy, except T3 (ML) and T6 (LL), which were identified at around 70% accuracy. None of the six tones produced by typically-developing children were perceived by the judges with adult-like accuracy, suggesting that three-year-old children had not mastered production of any of the six tones. More importantly, the acoustic findings showed that children's productions that were correctly identified by the judges by and large had the acoustic characteristics of adults' productions though not all acoustic parameters were adult-like. The acoustic characteristics of children's incorrect productions, however, were significantly different from adults' productions of the same tones but matched the expected acoustic characteristics of the (mis)perceived tones (e.g., children's incorrect T4 (LF) productions that were misperceived as T3 (ML) had significantly shallower slopes and higher pitch levels than children and adults' correct T4 (LF) productions). Similar acoustic findings were reported with children's Mandarin tones (Wong, 2012a). Taken together, the findings of these studies provided strong evidence that the perceptual ratings of the tones by the judges based on filtered stimuli is a valid and reliable method in determining tone production accuracy.

Cantonese Tone Perception in Children

Results of studies on children's Cantonese tone perception also raised questions about early Cantonese tone acquisition in children. Ching (1984) and Ciocca and Lui (2003) asked children to identify the six tones presented in monosyllabic words with the syllable /ji/ and reported that children did not correctly identify the six tones until 10;0. Considering that words formed by the syllable /ji/ may not be familiar to young children, Lee et al. (2015) examined children's Cantonese tone perception in familiar words. Two hundred typical Cantonese-speaking children and 25 adults listened to monosyllabic words and identified the tones in a picture-pointing task. The results indicated that children's tone perception improved between three to six years of age and did not reach adult-like accuracy until after

six years old (K. Lee, Chan, Lam, van Hasselt, & Tong, 2015). Combining the findings of these studies on tone perception with the findings on children's tone production in the studies presented above (e.g., So & Dodd, 1995; To et al., 2013), the results suggested that children mastered Cantonese tone production well before they correctly identified the tones, which contradicted the conventional assumption in speech acquisition that speech perception precedes speech production (Edwards, 1974).

Some studies showed that adults also made errors with Cantonese tone perception and production, suggesting that it is necessary to include an adult reference group to compare children's performance for determining tone mastery. For example, Ciocca and Lui (2003) found that adults made 5% - 20% errors when discriminating between T3 (ML) and T6 (LL) and between T2 (HR) and T5 (LR). Lee et al. (2015) reported that adults made 3% - 8% errors in monosyllabic tone identification. In Barry and Blamey (2004), the five adults made 12% errors when producing T2 (HR) and T5 (LR), while in Wong et al., (2017) the 19 adults made 20% - 30% errors in producing T3 (ML) and T6 (LL). The findings indicated that 100% accuracy may not be an appropriate criterion for making decision on children's tone mastery.

Relationship between Cantonese Tone Perception and Production in Children

The relation between children's tone production and perception ability has been inferred from studies that examined different groups of children. Wong et al., (2017) was the only study that examined Cantonese tone production and perception in the same group of children. Tone perception performance of the three-year-old children and their mothers was evaluated by a standardized tone identification test –Hong Kong Cantonese Tone Identification Test (CanTIT) (K. Y. S. Lee, 2012). The results showed that though children identified the six tones with significantly lower accuracy than adults, their tone perception accuracy (range = 72% - 92%) was higher than their tone production accuracy (range = 38% - 74%). Because

the study involved only one age group of children (i.e. 3-year-olds), it is unclear how children developed tone production and perception with time, when children acquire adult-like tone production and perception accuracy, and whether children achieve adult-like accuracy in tone perception before tone production. Therefore, further study should be undertaken to investigate children's development of Cantonese tone production and perception and the relation between Cantonese lexical tone production and perception.

To fill the research gap, this study adopted the methods used in Wong et al., (2017) to examine tone production and perception in 4 to 6-year-old typically-developing Cantonese-speaking children. The study is significant in providing information on the developmental trend of Cantonese tone acquisition which is important for providing clinical guidelines for evaluating and assisting children with tone difficulties. The study would also test the hypotheses in models of speech acquisition. If the prevailing assumption that children master suprasegmental features far earlier than segmental features is true, children would have acquired both tone perception and production well before six years of age, an age at which Cantonese-speaking children fully master all the Cantonese consonants and vowels (So & Dodd, 1995; To et al., 2013). In addition, if children acquire tone production before tone perception as previous studies have suggested, tone production accuracy would be significantly higher than tone perception accuracy in the same group of children. The specific research questions included: (1) What is the tone production ability of 4 to 6-year-old children? (2) What is the tone perception ability of 4 to 6-year-old children? (3) What is the developmental trend of tone production and perception ability in 4 to 6-year-old children? (4) Is there any relationship between children's tone production and perception?

Method

The present study followed the methods in Wong (2012a) and Wong, et al. (2017). Procedures were approved by the Human Research Ethics Committee at the University of Hong Kong.

Participants

Children. Fifty-one four- to six-year-old Cantonese-speaking children raised in Hong Kong with unremarkable developmental history were recruited (Appendix A). Eighteen of them were 4-year-olds (range = 4;0 - 4;10), eighteen were 5-year-olds (range = 5;0 - 5;10) and fifteen were 6-year-olds (range = 6;0 - 6;11). All children passed hearing screening at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz bilaterally at 20 dB using standard pure tone audiometry. The children were administered age-appropriate language tests. All 4-year-olds and 5-year-olds scored within normal limits in a language test -- Cantonese Oral Language Deficiency Early Identification Test for Pre-primary Children (學前兒童粵語表達能力識別測驗) (Po Leung Kuk, 2012), except one 4-year-old (C22) and one 5-year-old (C50), who were both excluded. All 6-year-olds scored within normal limits in another language test-- Cantonese Expressive Language Scales (粵語（香港）語言表達量表) (Hong Kong Education and Manpower Bureau, 2006), which was designed for children aged six years or above. All children used Cantonese as their home language and did not speak another tone language, except one 4-year-old child (C01), who was excluded because of his exposure to Mandarin at home. Thus, there were a total of 48 children in the study. Most of them came from families with high social economic status (Appendix A).

Adults. Twenty-nine native Cantonese-speaking mothers of the children (age between 27 to 37 years) who accompanied the children to the testing session and voluntarily participated in this study were recruited in the reference group. Mothers were adopted because they were more representative of the children's Cantonese tone input. One mother

(M01) whose first language was Mandarin was excluded. The twenty-eight mothers included in this study reported Cantonese as their first and primary home language (Appendix A).

Stimuli

To control for coarticulation and utterance prosodic effect on tones (Ma, Ciocca, & Whitehill, 2006; Wong & Strange, 2017), only monosyllabic words were adopted.

Tone production. Thirty monosyllabic Cantonese words (5 words x 6 tones) were represented in pictures. Twenty-four of the words were produced by over 80% of 30-month old children growing up in Hong Kong, as reported by parents in the Cantonese Communicative Developmental Inventory (CCDI) (Tardif, Fletcher, Zhang, Liang, & Zuo, 2008). To increase the number of words that would be tested in both perception and production, six words that were produced by fewer than 80% of 30-month-old children, but appeared in the tone perception test and were judged to be familiar to four- to six-year-old children were added (Appendix B).

Tone perception. Tone perception accuracy of participants was examined using 30 monosyllabic target words in CanTIT Form A (K. Y. S. Lee, 2012). The stimuli covered all 15 tonal contrasts of the six tones. Four pictures were presented in each trial, one representing the target word, another representing a word that formed a minimal tone pair with the target word (a tone distractor), and the other two representing vowel and consonant distractors.

Not all words in tone production were tested in tone perception (Appendix B). Only a subset of stimuli that formed minimal pair differed in tone with another word, met the selection criteria for familiar words in this study and was usually produced by children in monosyllabic context was tested in both perception and production.

Procedures

Each child participant attended a 1.5- to 2-hour session in a quiet room at home or in the University of Hong Kong. Before testing, mothers gave written informed consent and

filled out a questionnaire about the developmental history of the children, and provided the demographic information and language background of themselves and their children. Then tone production test was carried out before tone perception test to avoid delayed imitation. In the tone production test, child participants named randomized pictures representing 30 target monosyllabic words upon prompting questions (e.g. '呢個係咩嚟?' What is this?) and sentence completion cues (e.g. '我開…(門)' 'I open the…(door)') one by one. Five practice trials were given prior to the test to facilitate participants' understanding of the task. No target words were used in the prompts and no feedback was given. Their productions were audio-recorded for tone judgment.

After the picture naming task, the tone perception test, CanTIT, was carried out. There were 3 practice trials and 30 testing trials. Child participants listened to recordings of target words presented after a carrier phrase '邊幅係…' ('Which picture is …') over headphones and pointed to one of the four pictures presented on the screen. The experimenter input the responses on the computer. Mothers who volunteered to participate in the study were then invited to take the tone production and perception tests following the same procedures. Hearing and language screening for children were carried out at the end of the session.

Perceptual Judgment of tones produced by the participants

Judges. Five 18- to 23-year old native Cantonese-speaking undergraduate students at the University of Hong Kong, who were naïve to the design and hypotheses of the experiments, judged the tones produced by the participants in filtered stimuli. All judges were native to Hong Kong and used Cantonese as their primary language. No hearing, speech or language impairment was reported. They reported to use Cantonese over 80% of the time in daily life. Prior to tone ratings, all of them passed a tone screening test on filtered speech with over 90% accuracy.

Stimuli. 1440 child productions and 840 mother productions were low-pass filtered at 500 Hz and at 400 Hz, respectively (Wong, 2012b; Wong, Fu, & Cheung, 2017). Child productions were filtered at a higher cut-off frequency due to their higher mean F0. All productions were blocked by speakers and normalized to 68 dB for tone identification. The 77 blocks of stimuli were put in 20 experiments. Each experiment had up to four blocks and consisted of at least one child and one mother block. Each mother participant was paired up with her own child in separate blocks within the same experiment. Each experiment was around 30 minutes long.

Procedures. Judges listened to the filtered stimuli presented by a computer program over headphones in a quiet room at home or in the University of Hong Kong and identified the tones by inputting numbers representing the six tones on the screen. Blocks of stimuli and trials within blocks were presented randomly. Judges re-rated seven blocks of child productions and five blocks of adult productions (16% of the total productions) for intra-rater reliability.

Results

In the following analyses, first, tone production accuracy was presented, followed by tone perception accuracy, and then followed by correlation analyses.

Tone Production

Interjudge and intrajudge Reliability

The five judges were highly reliable in their tone judgments. Fleiss Kappa (k) was used to determine the degree of agreement among the five judges on their ratings of the tones produced by the participants. The results showed that the judges as a group reached substantial agreement on their tone ratings of adults' productions (k = 0.700) and children's and adults' productions combined (k = 0.613). The reliability of their ratings on children's productions was slightly lower, reaching moderate agreement, (k = 0.562).

Cohen's Kappa (k) was computed to examine the level of agreement in tone ratings in each pair of judges. All judges reached substantial agreement (k = 0.667 – 0.738) with other judges on rating adults' productions. Except one pair of judges who reached moderate agreement (k = 0.583), all other pairs of judges reached substantial agreement (k = 0.607 - 0.634) on rating children's and adults' productions combined. The judges were also highly reliable in their own tone ratings. Cohen's Kappa showed substantial intra-judge reliability in the five judges (k = 0.663- 0.732).

Tone Production Accuracy and Error Patterns of Adults

Tone production accuracy was defined as the percent of judges who correctly identified the target tones. Major error patterns were defined as substitution errors that constituted more than 10% of the trials. Tables 1-4 show the confusion matrices of the tones produced by each age group. Tone productions that were correctly identified by the judges were on the diagonals, while tones that were misperceived by the judges were off the diagonals. Major confusion patterns (i.e., more than 10% substitutions) were highlighted with light grey.

As shown in Table 1, mothers did not produce all tones with perfect accuracy. Most tones were perceived by the judges with higher than 80% accuracy, except T3 (ML) (63%) and T6 (LL) (74%). There was some confusion in the mothers' three level tones. Thirty-one percent of T3 (ML) was perceived as T6 (LL); 20% of T6 (LL) was perceived as T3 (ML), and 16% of T1 (HL) was identified as T3 (ML).

Tone Production Accuracy and Error Patterns of Children

Tables 2-4 show children's tone production accuracy and error patterns. The perceived accuracy of children's tones ranged from 39% to 80%, all lower than mothers'. The three groups of children shared similar accuracy rates, except that 6-year-olds (C6) had higher accuracy rates in T1 (HL) (54%) and T4 (LF) (80%). Children displayed more substitution

errors than adults and the three groups of children shared similar major error patterns, except that C6 had higher accuracy rates in T1 (HL) (54%) and T4 (LF) (80%). Overall, children tended to confuse among the three level tones, between the two rising tones, and between the Low-Falling and Low-Level tones in their productions. These error patterns were also found in adults' productions but with substantially fewer errors.

Tables 1-4

Order of Production Accuracy of Tones

A two-way mixed ANOVA was conducted using tones (T1 to T6) as within-subject factor, age groups (adults, C4, C5, and C6) as between-subject factor, and tone production accuracy as dependent variable to examine whether significant differences in tone production accuracy existed among tones and age groups. The results showed a significant main effect of age groups, $F(3, 72) = 37.555$, $p < .001$, $r = 0.78$, a significant main effect of tones, $F(3.867, 278.449) = 33.850$, $p < .001$, $r = 0.57$, and a significant interaction effect between tones and age groups on tone production accuracy, $F(11.602, 278.449) = 2.320$, $p = .008$, $r = 0.30$.

Posthoc pairwise comparisons were performed to examine order of accuracy of the six tones within each age group to determine whether adults and children produced any of the tones better than others. The results with Bonferroni adjustments for multiple comparisons were presented on the left in Appendix C. Adults produced T3 (ML) and T6 (LL) with significantly lower accuracy. For all child groups, T4 (LF), T5 (LR) and T2 (HR) were produced with higher accuracy than T1 (HL) and T3 (ML).

Differences in Tone Production Accuracy among the Age Groups

To determine which tones were produced by children with adult-like accuracy, pairwise comparisons with Bonferroni adjustments for multiple comparisons were carried out to compare production accuracy between adults and children for each tone and with all tones combined. As shown in Table 5, four-year-olds (C4) did not produce any tone with adult-like

accuracy. Five-year-olds (C5) produced T5 (LR) and T6 (LL) as accurately as adults, while C6 produced T4 (LF), T5 (LR) and T6 (LL) with accuracy rates comparable to adults. With all tones collapsed, no child group reached adult-like tone production accuracy.

Table 5

Tone Perception

Tone Perception Accuracy and Error Patterns in Adults and Children

Tone perception performance was examined based on participants' accuracy in identifying the target pictures on CanTIT (Form A), a standardized Cantonese tone perception test. Tone perception accuracies of all age groups were presented in Tables 6-9. Mothers perceived all tones with perfect accuracy (range = 99% - 100%) (Table 6). Children made more errors than adults (Tables 7-9). C5 and C6 perceived all tones with higher than 80% accuracy (Tables 8-9). T1 (HL) and T5 (LR) were perceived with the highest (96% - 99%) and the lowest (81% - 84%) accuracy, respectively. C4 appeared to attain lower accuracy than C5 and C6 and perceived three tones--T3 (ML), T5 (LR), and T4 (LF)--with lower than 80%. C4 also exhibited more error patterns than C5 and C6 (Tables 7-9). The major confusion pattern in C5 and C6 involved the perception of T5 (LR) as T2 (HR) (Tables 8-9). C4 displayed additional major error patterns, which involved bidirectional confusion of T2 (HR) and T4 (LF) (Table 7).

Tables 7-9

Order of Perception Accuracy of Tones

Because the data violated the assumptions of normality and sphericity for parametric statistics, Friedman tests were used to examine whether there were any differences in the accuracy rates among the six tones in each age group. Adults perceived the six tones with comparable accuracy, $\chi^2(5) = 4.000$, $p = 0.549$. However, C4, C5 and C6 perceived different

tones with different accuracies, $\chi^2(5) = 12.415$, $p = 0.03$; $\chi^2(5) = 16.697$, $p = 0.005$; and χ^2 (5) = 19.383, $p = 0.002$, respectively. Wilcoxon signed-rank tests were used to examine the order of accuracy of tones within each age group. The results after adjusting for multiple comparisons were shown on the right in Appendix C. Though there was a trend for T1 (HL) and T6 (LL) to be perceived with the highest accuracy in all child groups, no significant difference between the perceptual accuracy of the tones was found. In essence children perceived the six tones comparably.

Differences in Tone Perception Accuracy among the Age Groups

To determine whether children perceived any of the tones with adult-like accuracy, Kruskal-Wallis test was used to determine if perception accuracy of the tones in the child-groups was different from that of adults. The results showed a significant effect of age groups on tone perception accuracy, χ^2 (3) = 13.216, 27.323, 31.917, 26.689, 26.826, 23.375, and 57.225, for each of the six tones and all tones collapsed, $p = 0.004$ for T1 (HL) and $p < .001$ for all other comparisons. Mann-Whitney tests were conducted to examine differences in perception accuracy among the adult and child groups. As shown in Table 10, all child groups did not perceive any of the tones with adult-like accuracy, except for T1 (HL) in C6.

Table 10

Relationship between Tone Perception and Production Accuracy in Children

Development of tone production accuracy

Table 11 summarizes Pearson's r and the coefficients of determination (R^2) of the relations between children's tone production accuracy and age. As shown, children showed little improvement of tone production accuracy with age, R^2 for the six tones ranged from 0.001 to 0.054. When all tones were combined, there was a significant but small association between children's production accuracy and age. The top left panel in Appendix D shows the

scatterplot of children's overall tone production accuracy by age and Appendix E shows the development of tone production accuracy by each tone.

Table 11

Development of tone perception accuracy

Table 12 shows results of relations between children's tone perception accuracy and age. As presented, four- to six-year-old children showed significant improvement in perception accuracy of each tone with age, except for T5 (LR). R^2 ranged from 0.06 to 0.27. T3 (ML) demonstrated the largest development. With all tones combined, children's tone accuracy was significantly related to age with $R^2 = 0.335$, a large effect size. The top right panel in Appendix D shows the scatterplot of children's overall tone perception accuracy by age.

Table 12

Association between tone production and tone perception accuracy

Pearson correlation was performed on children's tone production accuracy of the 30 target words and tone perception accuracy in CanTIT Form A to determine how well children's tone perception ability predicted their production ability. Table 13 summarizes the correlation coefficients of tone production and perception accuracy by tone and with all tones combined. The bottom left panel in Appendix D shows the scatterplot of the relation. The results showed that although non-significant relations were found between tone production and perception accuracy for each tone, there was a significant relation with medium effect between tone production and perception accuracy when all tones were combined.

Table 13

To further analyze the relation between tone production and perception accuracy in the same set of words, a subset of 20 words tested in both tone production and perception was

selected. Table 14 and the bottom right panel in Appendix D show the correlation coefficients and the scatterplot of the relation, respectively. The results were similar to the comparison using the full set of production stimuli.

Table 14

Discussion

Cantonese Tone Production Accuracy in Adults

Consistent with the findings in previous studies, Cantonese adults did not produce all tones with ceiling accuracy. They did not always make clear distinction between T3 (ML) and T6 (LL). Confusions of the same tones in adults' speech have been reported in various studies (e.g., Wong et al., 2017; Barry & Blamey, 2004). Some Cantonese-speakers merge these tone categories (Mok, Zuo, & Wong, 2013). However, there was little evidence that the mothers included in this study were tone mergers. First, they all met the inclusion criteria for their family background and Cantonese-speaking proficiency. Second, none of the mothers were outliers in the tones with the lowest accuracy. Third, all mothers perceived all tones with high accuracy (99% - 100%) and no mothers produced any of the tones at 0% accuracy, suggesting that the mothers discriminated the tones in their perception and production. Therefore, the production accuracy of the 28 mothers likely represented natural variations in Cantonese tone production.

Cantonese Tone Production Accuracy in Children

Though children produced most of the tones with lower accuracy than adults, adult-like productions started to emerge at five years old. Wong et al., (2017) reported that three-year-old children produced none of the tones with adult-like accuracy. This study found that four-year-old children also did not produce any of the tones as accurately as adults; five-year-old and six-year-old children produced two tones (T5 (LR) and T6 (LL)) and three tones (T4 (LF), T5 (LR) and T6 (LL)), respectively, with adult-like accuracy. The combined

findings of Wong et al., (2017) and this study suggested that children started to master the production of some of the tones at five years of age.

Children demonstrated substantially more error patterns than mothers but the number of error patterns decreased with age. Mothers' major tone production errors involved confusions among the three level tones. Some bidirectional confusions between T2 (HR) and T5 (LL) were also observed though the confusions did not reach the criteria for major error patterns. Children showed all the three major error patterns found in mothers and additional confusions between T2 (HR) - T5 (LR) and T4 (LF) - T6 (LL). The number of error patterns decreased with age. Four-, five- and six-year-old children displayed eight, seven, and six major error patterns, respectively. All major error patterns found in the older age groups were also found in the younger age groups. Though some improvement was noted in the number of error patterns, children's tone production skills developed slowly with age, as indicated by the small correlation coefficients of tone production accuracy and age

Comparing the accuracy rates of the tones produced by children, the falling tone and the rising tones appeared to be easier for children than the three level tones. Wong et al., (2017), Barry and Blamey (2004) and Tse (1978) also found T4 (LF) to be an easy tone for Cantonese-speaking children. Coincidentally, studies that examined children's acquisition of Mandarin tones also reported higher accuracy in the falling tone (Wong, 2012a, 2012b, 2013). Wong (2012b) and Wong and Strange (2017) provided a physiological account for the order of acquisition of tones. It was hypothesized that the falling tone was produced mostly by passive relaxation rather than active control of the laryngeal muscles, and was, therefore, easier for children to master (Wong, 2012b).

The finding in this study that four- to six-year-old children had not produced all tones in monosyllabic words with adult-like accuracy was not consistent with the findings in previous studies that Cantonese-speaking children produced all tones in different contexts

correctly before the age of three years (So & Dodd, 1995; To et al., 2013; J. K.-P. Tse, 1978). The main reason for the difference may be due to the use of different methodology. In this study, judges identified tones in filtered stimuli from monosyllabic words. The method was evidenced to be accurate as tones correctly identified in filtered stimuli in adults had the acoustic properties of the target tones while tones incorrectly identified had acoustic properties different from the target tones and displayed acoustic properties of the tones selected by the raters (Wong, 2012b; Wong et al., 2017). The finding of late acquisition of Cantonese tones in this study agreed with the finding in Barry and Blamey (2004), which also controlled lexical biases in tone judgment by employing a non-native judge. They concluded that tone production was not yet mastered before six years of age. The slow mastery of Cantonese tone production found in this study also agreed with the findings in studies adopting similar methodology with Mandarin tones (e.g., Wong, 2013; Wong & Strange, 2017).

Cantonese Tone Perception Accuracy in Children

Mothers perceived all the tones with perfect accuracy. Children from four- to six years old perceived all the tones with higher than 80% accuracy, much higher than their production accuracy and higher than the perception accuracy of 3-year-olds, as reported in Wong et al, (2017). Despite the high perceptual accuracy, none of the tones were perceived by children with adult-like accuracy, except for T1 (HL) in six years olds, indicating that mastery of Cantonese tone perception did not emerge until six years of age. The findings were comparable with studies that reported that children perceived tones at adult-like accuracy after six years old (K. Lee et al., 2015) at around ten years of age (Ciocca & Lui, 2003). T1 (HL) appeared to be easier for children to identify because there was a trend for T1 (HL) to be perceived with higher accuracy rates in four- to five-year-olds. Wong et al., (2017) also found significantly higher perception accuracy of T1 (HL) in 3-year-old children. This

may be due to the larger pitch differences between T1 (HL) and the other tones (Figure 1, Wong & Chan, 2018), making it more salient than others.

The number of confusion patterns in children's tone perception decreased as they got older. Four-year-old children displayed the same two major substitution patterns found in three-year-old children in Wong et al., (2017), while five- and six-year-old children had only one major error pattern. Taken together, three and four-year-old children had more difficulty discriminating between T2 (HR) and T4 (LF) and between the two rising tones (i.e., T2 (HR) and T5 (LR)), while five and six-year-old children had difficulty discriminating the two rising tones, only. Perceptual difficulty of the two rising tones has also been reported in Lee et al. (2015) with 3- to 6-year-old children and could be explained by the similar F0 shapes and F0 onsets of the two rising tones (Lee et al., 2015; Wong & Chan, 2018).

Children made moderate to substantial improvement in tone perception accuracy from four- to six years old, as evidenced by the correlations between children's perception accuracy and age. The developmental trends reported in this study aligned with the findings of gradual increase in perception accuracy from three to six years in two studies (Ciocca & Lui, 2003; K. Lee et al., 2015).

As shown, tone perception accuracy had not been fully developed in four- to six-year-old children. Together with the finding in Wong et al., 2017 that three-year-old children did not identify any of the tones with adult-like accuracy, it appeared that children gradually improved their tone perception ability between three to six years old. When they reached the age of six years, adult-like tone perception started to emerge.

Relationship between tone production and perception in Children

Previous findings in tone perception and production in separate groups of children suggested that Cantonese children accurately produced the six tones three or more years before they correctly identified the tones in monosyllabic words. This study and Wong et al.,

(2017) examined Cantonese tone perception and production in the same groups of children. The results did not support the claim that children acquire tone production before tone perception, and suggested the contrary. Children have not fully mastered Cantonese tone production or perception at the age of six years. However, they achieved higher accuracy and showed more improvement in tone perception than production.

Children's tone perception ability did not predict their tone production ability, as indicated by the weak correlations between children's tone production and perception accuracy. In addition, children perceived all the six tones with comparable accuracy, but produced them with significantly different accuracy. Moreover, error patterns in tone production were not found in tone perception. Furthermore, all tones, except T4 (LF), showed only small and non-significant correlations between perception and production even when correlations were performed on the same set of familiar words. These findings suggested that other factors contributed to the slow development in children's tone production.

One factor that contributed to the slow mastery of tone production could be immature speech motor control in children. It was reported that children's speech motor skills developed gradually. Children continued to refine their speech motor control and optimize their speech coordination from six years old to adolescence (Green, Moore, Higashikawa, & Steeve, 2000) and did not master mature speech motor control until after sixteen years old (Smith, 2006). This physiological account was supported by Wong (2013) who found that Mandarin tones with more articulatory complexity, namely the dipping tone, was acquired the last in three to five-year-old children and by Wong and Strange (2017) who reported that children produced the same Mandarin tones with lower accuracy when they occurred in disyllabic tone combinations that formed more complex F0 contours (Wong, 2013; Wong & Strange, 2017). Another possible factor could be inconsistent tone input in daily life. Typical adults and tone mergers have been reported to confuse or fail to make the contrasts between

the two rising tones, the two lower level tones, and T4 (LF) and T6 (LL) in their productions (Bauer, Kwan-hin, & Pak-man, 2003; Mok et al., 2013), making it harder for children to form exact phonological representations necessary for correct tone productions.

Conclusions

Overall, this study found that four- to six-year-old children have not mastered perception or production of the six Cantonese tones in familiar monosyllabic words. . Cantonese-speaking children demonstrated substantial tone perception development but slow tone production development between the ages of four to six years. Children's tone perception ability does not predict their tone production ability. Adult-like tone production and perception ability started to emerge at five and six years old.

Theoretical and Clinical Implications

Several findings in the current study may shed light on models and theories of speech development in children. First, the results in this study contradict the long-held assumption on speech acquisition that lexical tones are acquired rapidly and early before children fully master the segmental system in their language (So & Dodd, 1995; To et al., 2013; J. K.-P. Tse, 1978). As most Cantonese children have acquired the segmental features in Cantonese by the age of six years (So & Dodd, 1995; To et al., 2013), the continual development of tones after six years of age, suggested in this study, does not support the claim that children acquire supra-segmental features much earlier than segmental features.

Second, the findings of this study do not support the observation in previous studies on children's Cantonese tone acquisition that children master tone production well before tone perception, and support the prevalent assumption in speech development models that speech perception precedes speech production. While previous studies reported that children acquired tone production before the age of three years and mastered tone perception at around six to ten years (Ching, 1984; K. Lee et al., 2015; So & Dodd, 1995; To et al., 2013),

the results of this study show that pre-school children's tone production ability fall behind their tone perception ability in terms of accuracy rates, number of error patterns and rates of development.

Third, the weak relationship between children's production and perception ability does not support the previous assumption that speech perception ability highly predicts speech production ability (Lotto, Hickok, & Holt, 2009). Other factors such as speech motor control and inconsistent tone input may affect children's tone production development.

With respect to clinical implications, tone perception and production difficulties have been reported in different clinical populations, such as children with poor comprehension (Zhang et al., 2014), dyslexia (Li & Ho, 2011), and hearing loss (Ciocca et al., 2002). This study, together with Wong et al., (2017), provides detailed information on the development of Cantonese tone perception and production in pre-school children. The order of acquisition of tones and the error patterns reported can serve as reference in assessing and treating pre-school children with tone difficulties and for future research on tone development in different clinical populations. Because children as old as six years of age have not mastered the perception and production of the six tones, further research utilizing the same methodology is needed to track the development of tone perception and production in older children and to determine the age of acquisition of Cantonese tones.

References

Barry, J. G., & Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *The Journal of the Acoustical Society of America, 116*(3), 1739-1748. doi:10.1121/1.1779272

Bauer, R. S., Kwan-hin, C., & Pak-man, C. (2003). Variation and merger of the rising tones in Hong Kong Cantonese. *Language Variation and Change, 15*(2), 211-225. doi:10.1017/S0954394503152039

Chen, L.-M., Ni, H.-C., Kuo, T.-W., & Hsu, K.-L. (2012). *Proceedings of the Twenty-Fourth Conference on Computational Linguistics and Speech Processing.* Paper presented at the Conference on Computational Linguistics and Speech Processing, Taiwan.

Cheung, H., Chung, K. K. H., Wong, S. W. L., McBride-Chang, C., Penney, T. B., & Ho, C. S. H. (2009). Perception of tone and aspiration contrasts in Chinese children with dyslexia. *Journal of Child Psychology and Psychiatry, 50*(6), 726-733. doi:10.1111/j.1469-7610.2008.02001.x

Ching, T. Y. C. (1984). Lexical tone pattern learning in Cantonese children. *Language Learning and Communication, 3*(3), 243-414.

Ciocca, V., Francis, A. L., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *The Journal of the Acoustical Society of America, 111*(5), 2250-2256. doi:10.1121/1.1471897

Ciocca, V., & Lui, J. (2003). The development of the perception of Cantonese lexical tones. *Journal of Multilingual Communication Disorders, 1*(2), 141-147. doi:10.1080/1476967031000090971

Clumeck, H. (1980). The acquisition of tone. *Child phonology, 1*, 257-275.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research, 229*(1), 132-147. doi:http://dx.doi.org/10.1016/j.heares.2007.01.014

Dehaene-Lambertz, G., & Houston, D. (1998). Faster Orientation Latencies Toward Native Language in Two-Month-Old Infants. *Language and Speech, 41*(1), 21-43. doi:10.1177/002383099804100102

Edwards, M. L. (1974). Perception and production in child phonology: the testing of four hypotheses. *Journal of child language, 1*(2), 205-219. doi:10.1017/S0305000900000659

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The Physiologic Development of Speech Motor ControlLip and Jaw Coordination. *Journal of Speech, Language, and Hearing Research, 43*(1), 239-255. doi:10.1044/jslhr.4301.239

Hua, Z. (2002). *Phonological development in specific contexts: Studies of Chinese-speaking children* (Vol. 3): Multilingual Matters.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of child language, 27*(01), 3-42.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608.

Lee, K., Chan, K. T. Y., Lam, J. H. S., van Hasselt, C. A., & Tong, M. C. F. (2015). Lexical tone perception in native speakers of Cantonese. *International Journal of Speech-Language Pathology, 17*(1), 53-62. doi:10.3109/17549507.2014.898096

Lee, K. Y. S. (2012). *The Cantonese Tone Identification Test (CanTIT)*

Hong Kong: Department of Otorhinolaryngology, Head & Neck Surgery, The Chinese University of Hong Kong.
Li, W.-S., & Ho, C. S.-H. (2011). Lexical tone awareness among Chinese children with developmental dyslexia*. *Journal of child language, 38*(4), 793-808. doi:10.1017/S0305000910000346
Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences, 13*(3), 110-114. doi:https://doi.org/10.1016/j.tics.2008.11.008
Ma, J. K. Y., Ciocca, V., & Whitehill, T. L. (2006). Effect of intonation on Cantonese lexical tones. *The Journal of the Acoustical Society of America, 120*(6), 3978-3987. doi:10.1121/1.2363927
Mok, P. P. K., Zuo, D., & Wong, P. W. Y. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change, 25*(3), 341-370. doi:10.1017/S0954394513000161
Oller, D. K., & Eilers, R. E. (1975). Phonetic Expectation and Transcription Validity. *Phonetica, 31*(3-4), 288-304.
Simons, G. F., & Charles, D. F. (2017). *Ethnologue: Languages of the World. Online version* (20 ed.): SIL International.
Smith, A. (2006). Speech motor development: Integrating muscles, movements, and linguistic units. *Journal of communication disorders, 39*(5), 331-349.
So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of child language, 22*(3), 473-495. doi:10.1017/S0305000900009922
Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Zuo, Q. H. (2008). *The Chinese Communicative Development Inventory (Putonghua and Cantonese versions): Manual, Forms, and Norms.* . Peking University Medical Press. .
To, C. K. S., Cheung, P. S. P., & McLeod, S. (2013). A Population Study of Children's Acquisition of Hong Kong Cantonese Consonants, Vowels, and Tones. *Journal of Speech, Language, and Hearing Research, 56*(1), 103-122. doi:10.1044/1092-4388(2012/11-0080)
Tse, A. C.-Y. (1992). *The acquisition of Cantonese phonology-a case study.* (M. Phil. M. Phil. Thesis), The University of Hong Kong, Hong Kong.
Tse, J. K.-P. (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of child language, 5*(2), 191-204. doi:10.1017/S0305000900007418
Tuaycharoen, P. (1977). *The Phonetic and Phonological Development of a Thai Baby: From Early Communicative Interaction to Speech* (Doctoral Doctoral), University of London.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development, 7*(1), 49-63.
Wong, P. (2012a). Acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions. *Journal of Phonetics, 40*(1), 141.
Wong, P. (2012b). Monosyllabic Mandarin tone productions by 3-year-olds growing up in Taiwan and in the United States: Interjudge reliability and perceptual results. *Journal of Speech, Language, and Hearing Research, 55*(5), 1423-1437.
Wong, P. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *The Journal of the Acoustical Society of America, 133*, 434-443. doi:10.1121/1.4768883

Wong, P., & Chan, H.-Y. (2018). Acoustic characteristics of highly distinguishable Cantonese tones. *The Journal of the Acoustical Society of America*.

Wong, P., Fu, W.-M., & Cheung, E. Y.-L. (2017). Cantonese-speaking children do not acquire tone perception before tone production -- A perceptual and acoustic study of three-year-olds' monosyllabic tones. *Frontiers in psychology*.

Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research, 48*(5), 1065-1079.

Wong, P., & Strange, W. (2017). Phonetic complexity affects children's Mandarin tone production accuracy in disyllabic words: A perceptual study. *PLoS One, 12*(8), e0182337. doi:10.1371/journal.pone.0182337

Yip, M. (2002). Tone. Cambridge, UK: Cambridge University Press.

Zhang, J., McBride-Chang, C., Wong, A. M. Y., Tardif, T., Shu, H., & Zhang, Y. (2014). Longitudinal correlates of reading comprehension difficulties in Chinese children. *Reading and Writing, 27*(3), 481-501. doi:10.1007/s11145-013-9453-4

Table 1. Confusion matrix of adults' tone productions

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	80	0	16	0	1	3
T2	0	93	0	0	7	0
T3	2	0	63	1	2	31
T4	0	0	1	91	1	7
T5	0	8	1	1	88	2
T6	0	0	20	3	2	74

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 2. Confusion matrix of four-year-olds' tone productions

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	46	1	38	1	2	13
T2	1	60	1	2	35	2
T3	8	0	42	4	4	42
T4	1	0	5	77	1	17
T5	1	20	3	3	65	9
T6	4	1	25	14	4	52

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 3. Confusion matrix of five-year-olds' tone productions

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	39	0	42	1	3	15
T2	2	64	0	2	30	1
T3	8	1	41	2	8	40
T4	2	0	7	66	3	22
T5	0	15	2	2	79	2
T6	5	0	22	4	3	65

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 4. Confusion matrix of six-year-olds' tone productions

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	54	0	35	0	2	9
T2	0	65	0	2	33	1
T3	3	1	44	3	5	45
T4	1	0	4	80	1	14
T5	0	20	1	1	76	2
T6	3	0	26	3	6	61

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 5. T*one production accuracy between children and adults*

Tone Production			
Target tones	**Adult vs. C4**	**Adult vs. C5**	**Adult vs. C6**
T1	Adult > C4**	Adult > C5**	Adult > C6**
T2	Adult > C4**	Adult > C5**	Adult > C6**
T3	Adult > C4**	Adult > C5**	Adult > C6**
T4	Adult > C4**	Adult > C5**	Adult = C6
T5	Adult > C4**	Adult = C5	Adult = C6
T6	Adult > C4**	Adult = C5	Adult = C6
All Tones	Adult > C4**	Adult > C5**	Adult > C6**

Note. "*" and "**" indicates statistical difference at .05 level and .01 level after Bonferroni corrections, respectively. ">" indicates "significantly higher than" and "<" indicates "significantly lower than". Non-significant comparisons are marked by shaded cells.

Table 6. Confusion matrix of adults' tone perception

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	100	0	0	0	0	0
T2	0	100	0	0	0	0
T3	0	0	100	0	0	0
T4	0	1	0	99	0	0
T5	0	1	0	0	99	0
T6	0	0	0	0	0	100

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 7. Confusion matrix of four-year-olds' tone perception

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	93	0	3	0	3	3
T2	1	80	3	13	1	3
T3	3	5	74	9	3	8
T4	0	16	3	78	3	1
T5	0	13	3	4	76	5
T6	8	1	1	1	5	84

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 8. Confusion matrix of five-year-olds' tone perception

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	96	0	2	0	1	0
T2	0	91	1	1	6	1
T3	1	1	86	5	0	7
T4	0	8	1	85	6	0
T5	0	11	4	1	81	4
T6	1	0	1	0	2	95

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 9. Confusion matrix of six-year-olds' tone perception

Target tones	Judges' responses (%)					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**
T1	99	0	1	0	0	0
T2	0	89	0	3	8	0
T3	0	0	93	0	0	7
T4	0	7	3	88	3	0
T5	0	13	0	0	84	3
T6	3	0	0	0	1	96

Note. Correct identifications are marked by dark shaded cells. Light shaded cells mark error patterns that constitute more than 10% of the trials.

Table 10. T*one perception accuracy between children and adults*

Tone Perception			
Target tones	**Adult vs. C4**	**Adult vs. C5**	**Adult vs. C6**
T1	Adult > C4**	Adult > C5*	Adult = C6
T2	Adult > C4**	Adult > C5**	Adult > C6**
T3	Adult > C4**	Adult > C5**	Adult > C6**
T4	Adult > C4**	Adult > C5**	Adult > C6**
T5	Adult > C4**	Adult > C5**	Adult > C6**
T6	Adult > C4**	Adult > C5**	Adult > C6*
All Tones	Adult > C4**	Adult > C5**	Adult > C6**

Note. "*" and "**" indicates statistical difference at .05 level and .01 level after Bonferroni corrections, respectively. ">" indicates "significantly higher than" and "<" indicates "significantly lower than". Non-significant comparisons are marked by shaded cells.

Table 11. Development of tone production

Development of Tone Production				
Target Tone	***r***	***R^2***	**Effect size**	***p***
T1	0.195	0.038	small	0.182
T2	0.148	0.022	small	0.32
T3	0.032	0.001	small	0.86
T4	0.077	0.006	small	0.611
T5	0.089	0.008	small	0.545
T6	0.232	0.054	small	0.111
All Tones	0.292	0.085	small	0.044*

Table 12. Development of tone perception

Development of Tone Perception				
Target Tone	***r***	***R^2***	**Effect size**	***p***
T1	0.379	0.144	medium	0.008**
T2	0.327	0.107	medium	0.024*
T3	0.520	0.270	large	<.001**
T4	0.336	0.113	medium	0.02*
T5	0.245	0.060	small	0.094
T6	0.395	0.156	medium	0.006**
All Tones	0.579	0.335	large	<.001**

Table 13. Relation between Tone Perception and Production (N=30[a])

Relation between Tone Perception and Production (N=30[a])				
Target Tone	***r***	***R^2***	**Effect size**	***p***
T1	-0.148	0.022	small	0.315
T2	0.064	0.004	small	0.667
T3	0.127	0.016	small	0.390
T4	-0.014	0.000	small	0.926
T5	0.090	0.008	small	0.543
T6	0.184	0.034	small	0.210
All Tones	0.44	0.194	medium	<0.01**

Note. [a] Tone production accuracy was based on the production of the 30 target words in the study and tone perception accuracy was based on the scores of CanTIT Form A.

Table 14. Relation between Tone Perception and Production (**N=20[a]**)

Relation between Tone Perception and Production in the Same Words (N=20[b])				
Target Tone	***r***	***R^2***	**Effect size**	***p***
T1	0.146	0.021	small	0.323
T2	0.207	0.043	small	0.157
T3	0.276	0.076	small	0.056
T4	0.388	0.151	medium	<0.01**
T5	0.134	0.018	small	0.362
T6	0.162	0.026	small	0.271
All Tones	0.381	0.145	medium	<0.01**

Note. [a] Tone production and perception accuracy was based on a subset of 20 words that were tested in both perception and production.

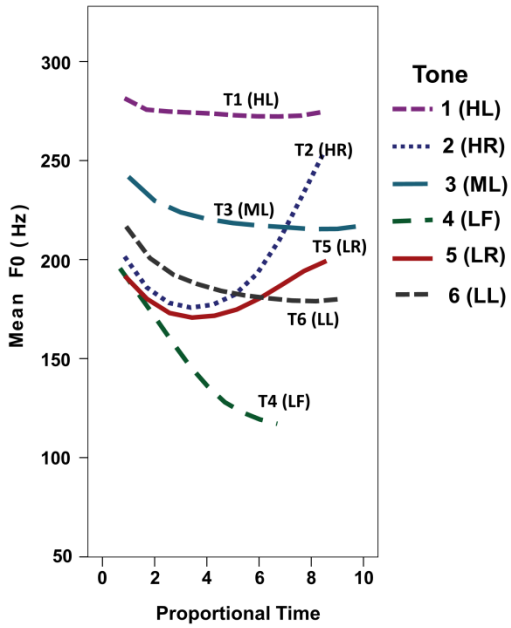


Figure 1. Fundamental frequency contours of Cantonese tones.
Notes. Each contour represents the mean F0 of 198 productions (6 syllables x 3 repetitions x 11 speakers).

Appendix A. Demographic Information of children in the study

Children's Information				Family Information					
Code	**Child Age**	**Age Group**	**Daily use (%) of Cantonese**	**Siblings[a]**	**Mother's Code**	**Mother's Age (Yrs)[c]**	**Mother's Education Level[d]**	**Father's Education Level[d]**	**Household Income[e]**
4-year-old children (C4)									
C02	4;2	C4	100	NA	M02	31	Mid	Mid	Mid
C06	4;10	C4	90	NA	M06	29	High	Mid	High
C08	4;0	C4	100	NA	-	-	Mid	Mid	High
C09	4;0	C4	100	NA	M09	31	High	High	Mid
C11	4;4	C4	100	NA	M11	28	Mid	High	Mid
C12	4;8	C4	100	NA	M12	37	High	High	High
C14	4;0	C4	90	NA	M14	35	High	High	High
C15	4;10	C4	100	NA	M15	33	Low	Low	Mid
C16	4;4	C4	95	C17	M16[b]	34	High	Mid	High
C18	4;7	C4	99	NA	-	-	Mid	High	High
C20	4;6	C4	90	NA	M20	27	High	High	High
C28	4;6	C4	90	NA	M28	36	High	High	High
C29	4;5	C4	100	NA	M29	37	Mid	High	Mid
C31	4;8	C4	98	NA	M31	33	High	High	Mid
C37	4;0	C4	95	C32	-	-	High	High	High
C44	4;7	C4	90	C45	M44[b]	33	High	High	High
5-year-old children (C5)									
C05	5;3	C5	100	NA	M05	30	Mid	Mid	Mid
C07	5;0	C5	100	NA	M07	30	Mid	High	High
C10	5;0	C5	100	NA	M10	31	Mid	High	High
C13	5;8	C5	100	NA	M13	36	High	High	High
C17	5;6	C5	95	C16	M16[b]	34	High	Mid	High
C19	5;3	C5	80	NA	-	-	High	High	Mid

C23	5;9	C5	100	NA	-	-	High	High	High
C24	5;10	C5	90	NA	M24	32	High	High	Mid
C25	5;4	C5	95	NA	-	-	Mid	Mid	Mid
C26	5;7	C5	100	C27	-	-	High	High	High
C27	5;7	C5	90	C26	-	-	High	High	High
C30	5;3	C5	90	NA	M30	36	High	High	High
C36	5;2	C5	95	C35	-	-	High	High	High
C42	5;2	C5	100	C43	M42[b]	36	High	High	High
C47	5;6	C5	100	C46	M46[b]	36	High	High	Mid
C48	5;1	C5	95	NA	M48	35	High	High	High
C49	5;5	C5	90	NA	M49	34	High	Mid	High
6-year-old children (C6)									
C03	6;1	C6	95	NA	M03	35	High	Low	Mid
C04	6;10	C6	90	NA	M04	33	Mid	Mid	High
C21	6;0	C6	95	NA	-	-	Mid	Mid	High
C32	6;6	C6	90	C37	-	-	High	High	High
C33	6;8	C6	100	NA	-	-	Mid	Mid	Mid
C34	6;5	C6	80	NA	-	-	High	High	Mid
C35	6;8	C6	95	C36	-	-	High	High	High
C38	6;6	C6	100	NA	-	-	Mid	High	Mid
C39	6;1	C6	90	NA	-	-	Mid	Mid	High
C40	6;6	C6	100	NA	-	-	High	High	Mid
C41	6;11	C6	95	NA	M41	36	High	High	Mid
C43	6;10	C6	100	C42	M42[b]	36	High	High	High
C45	6;1	C6	90	C44	M44[b]	33	High	High	High
C46	6;11	C6	100	C47	M46[b]	36	High	High	Mid
C51	6;0	C6	100	NA	M51	36	Low	Low	Mid

Notes.

[a] "NA" means "Not Applicable". Relationships between siblings: C16 and C17 -- Sister, Brother; C26 and C27 -- Identical twins; C32 and C37 -- Brother, Sister; C35 and C36 -- Sister, Sister; C42 and C43 -- Sister, Sister; C44 and C45 -- Sister, Brother; C46 and C47 -- Sister, Brother

[b] Mother of 2 Siblings

[c] "-" means "missing information or no information"

[d] Classification of Education level: "Low" = Primary - Lower secondary, "Mid" = Upper secondary, "High" = Post-secondary

[e] Classification of household income: "Low" = income of the lowest 30% of the total domestic household, "Mid" = income of the mid 35% of the total domestic household. "High" = income of the highest 35% of the total domestic household based on the results of the 2016 Population By-census conducted by the Census and Statistics Department of Hong Kong.

Appendix B. Tone Production Test Stimuli

Tone	Chinese Word	English Meaning	IPA[a]	In CanTIT[b] Form 30A	Tested in Perception and Production	Familiarity %[c]
T1 (HL)	杯	cup	/pui1/	No	No	96
	書	book	/sy1/	Yes	Yes	94
	貓	cat	/mau1/	No	No	93
	煲	pot	/pou1/	Yes	Yes	56
	湯	soup	/thɔŋ1/	No	Yes	90
T2 (HR)	糖	candy	/thɔŋ2/	No	Yes	97
	帽	hat	/mou2/	No	Yes	91
	手	hand	/sɐu2/	No	Yes	99
	魚	fish	/jy2/	No	Yes	99
	梨	pear	/lei2/	Yes	Yes	46
T3 (ML)	菜	vegetable	/tshɔi3/	No	No	91
	喊	cry	/ham3/	No	No	93
	鏡	mirror	/kɛŋ3/	No	Yes	87
	瘦	thin	/sɐu3/	Yes	Yes	N/A
	褲	pants	/fu3/	No	Yes	96
T4 (LF)	鞋	shoe	/hai4/	No	Yes	99
	頭	head	/thɐu4/	No	No	96
	門	door	/mun4/	Yes	Yes	93
	毛	towel	/mou4/	No	Yes	84
	床	bed	/tshɔŋ4/	No	No	88
T5 (LR)	被	blanket	/phei5/	No	No	93
	雨	rain	/jy5/	No	Yes	94
	眼	eye	/ŋan5/	No	No	97
	馬	horse	/ma5/	Yes	Yes	90
	蟹	crab	/hai5/	Yes	Yes	50
T6 (LL)	飯	rice	/fan6/	No	No	93
	麵	noodle	/min6/	No	No	93
	鼻	nose	/pei6/	Yes	Yes	96
	樹	tree	/sy6/	Yes	Yes	79
	腩	tongue	/lei6/	Yes	Yes	62

[a] International Phonetic Alphabet

[b] Hong Kong Cantonese Tone Identification Test (CanTIT).

[c] Percent of 30-month-old Cantonese-speaking children growing up in Hong Kong who produced the words based on parents' reports in the Cantonese Communicative Development Inventory (CCDI) (Tardif, Fletcher, Liang, & Kaciroti, 2009).

HL=High-level; HR=High-rising; ML=Mid-level; MR= Mid-rising; LL=Low-level; LR= Low-rising

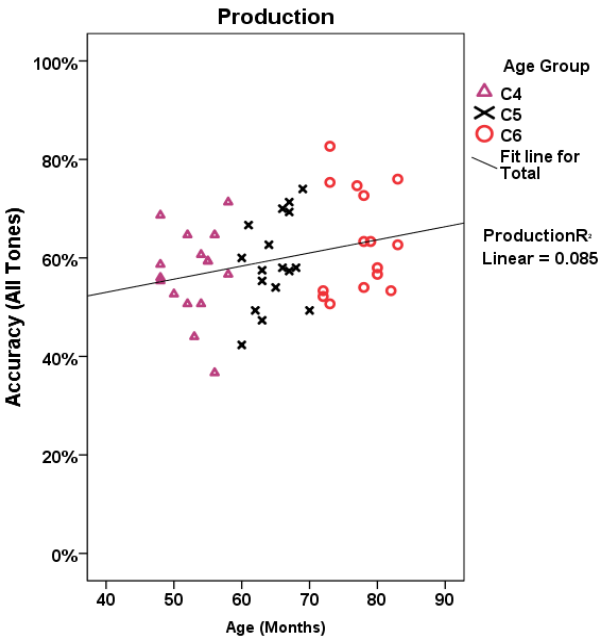
Appendix C. *Order of production and perception accuracy of the six tones by age groups*

Age Group	Order of production accuracy of the six tones	Order of perception accuracy of the six tones
	Tone Production	Tone Perception
Adult	T2 = T4 = T5 = T1 > (T1 =) T6 = T3	T1 = T2 = T3 = T6 = T4 = T5
C4	T4 = T5 = T2 > (T5 = T2 =) T6 = (T5 = T2 =) T1 = (T2 =) T3	T1 = T6 = T2 = T4 = T5 = T3
C5	T5 = T4 = T6 = T2 > T3 = T1	T1 = T6 = T2 = T3 = T4 = T5
C6	T4 = T5 = T2 > (T5 = T2 =) T6 = (T5 = T2 =) T1 > (T1=) T3	T1 = T6 = T3 = T2 = T4 = T5

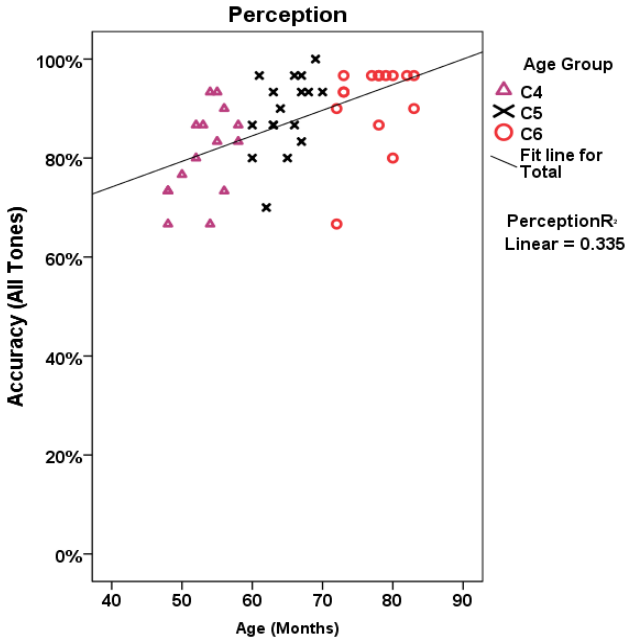
Note. "=" indicates "not significantly different from" and ">" indicates "significantly more accurate than". When two tones are significantly different (e.g., T1 > T6) and another tone (e.g., T3) is not significantly different from these two tones (i.e. T1 or T6), the relationship is notated as: T1 (= T3) > (T3 =) T6, meaning that T1 is significantly more accurate than T6, while T3 is not significantly different from either T1 or T6.

Appendix D. Children's tone production and perception development and association between children's tone production and perception

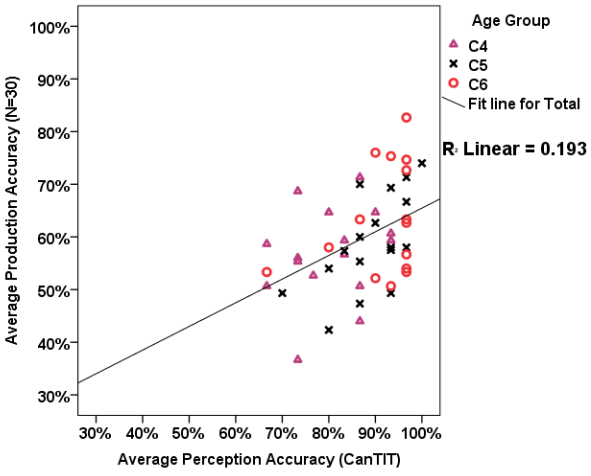
(a) Developmental Trend of Tone Production



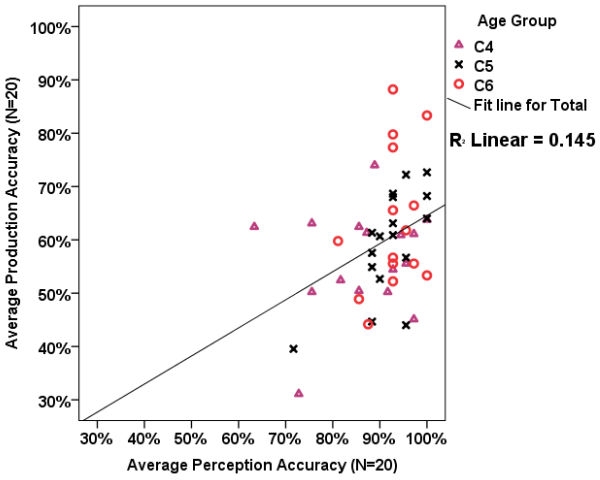
(c) Developmental Trend of Tone Perception



(b) Relation between CanTIT[a] Scores and Production Accuracy of the 30 Target Words



(d) Relation between Production and Perception Accuracy of the Same 20 Words



[a]Hong Kong Cantonese Tone Identification Test (CanTIT).

Appendix E. Development of the production of individual tones