Lexical tone perception and production in Cantonese-speaking children with childhood apraxia of speech: A pilot study

Eddy C. H. WONG[a], Min Ney WONG[a, b], Shelley L. VELLEMAN[c], Michael C. F. TONG[d] and Kathy Y. S. LEE[d]

[a]Department of Chinese and Bilingual Studies, the Hong Kong Polytechnic University, Hong Kong SAR, China; [b]Research Centre for Language, Cognition, and Neuroscience, the Hong Kong Polytechnic University, Hong Kong SAR, China; [c]Department of Communication Sciences and Disorders, the University of Vermont, Burlington, Vermont, USA; [d]Department of Otorhinolaryngology, Head and Neck Surgery & The Institute of Human Communicative Research, the Chinese University of Hong Kong, Hong Kong SAR, China.

Correspondence:

Eddy C. H. WONG

eddychwong@gmail.com

EF701, the Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China.

Source of support: This study was supported by the Office of Research and Knowledge Transfer Service at the Chinese University of Hong Kong and the Social Innovation and Entrepreneurship Development Fund (SIE Fund), which was granted to the first, fourth and last authors. The reference number is KPF18HLF26.

Word count: 6574 excluding the cover page and tables

Lexical tone perception and production in Cantonese-speaking children with childhood apraxia of speech: A pilot study

Introduction: Childhood apraxia of speech (CAS) is a paediatric motor speech disorder. We investigated the lexical tone perception and production abilities of children with CAS and the relationships between the two. **Methods:** Three children with CAS, aged between 3;7 and 5;8, were given the Cantonese Tone Identification Test (CANTIT) and the Hong Kong Cantonese Articulation Test (HKCAT) for assessment of tone perception and production, respectively. Accuracy and error patterns were investigated based on their performance on the two tests. Correlation analysis was performed on children's perception and production scores. **Results:** Two children scored at the lowest rank on the CANTIT, while one child obtained a Z score of 0. All children scored three standard deviations below the mean on the HKCAT. No statistical differences were found among the six tones with respect to perception accuracy, $H(5) = 3.731$, $p = 0.589$. Error analysis showed that children with CAS demonstrated more confusion on perceiving tones compared with TD peers. There were no main effects for task ($F(1,2) = 0.040$, $p = 0.859$) or tone ($F(5,10 = 0.997$, $p = 0.467$); nor were there task or tone interaction effects on perception versus production accuracy ($F(5,10) = 1.772$, $p = 0.206$). Tone perception and production accuracy were not significantly correlated ($r^2 = 0.181$, $p = 0.078$). **Discussion:** Tone perception deficits were evident in two out of three children with CAS, while all children had lexical tone production difficulties. In this small sample, tone production was more universally affected than tone perception.

Keywords: Lexical tone; perception; production; childhood apraxia of speech

Introduction

Childhood apraxia of speech (CAS) is a type of pediatric motor speech disorder (Shriberg et al., 2019) marked by impairment of motor planning and/or programming of speech movements (Iuzzini-Seigel et al., 2015). Speech production characteristics clearly dominate the clinical features in English speakers with CAS (Iuzzini-Seigel & Murray, 2017) as well as speakers of tonal languages with CAS (Wong et al., 2020). The American Speech-Language-Hearing Association [ASHA] (2007) lists three core features of CAS, including inconsistent errors, lengthened and disrupted coarticulatory transitions, and deficits in prosody such as lexical stress errors. These core features imply that CAS is a speech disorder that is characterized by a deficit of both segmental and suprasegmental skills.

Speech perception skills in children with CAS have been experimentally explored based on scientific support for relationships between speech perception and production in children with speech sound disorders (Berti et al., 2021; Cabbage et al., 2016). Past studies of speech perception skills at the segmental level in children with CAS can be divided into those that focus on lower-order auditory perception/discrimination skills (e.g., non-word discrimination) and those that require higher-order processing, such as categorization, lexical identification, and rhyming (Nijland, 2009). Some discrimination studies have shown that children with CAS are less proficient at speech discrimination than children who are typically developing (TD), including perceptual differentiation of vowel continua (Maassen et al., 2003), nonwords (Bridgeman & Snowling, 1988; Nijland, 2009), and real words (Groenen et al., 1996). However, Zuk et al. (2018) used a [da]-[ga] continuum to demonstrate that children with CAS and concomitant language impairment perform more poorly on discrimination than those with CAS only or TD, while children with CAS only show similar syllable discrimination thresholds as the TD group. These authors concluded that poor speech perception is not a core deficit of CAS.

These findings suggest that concomitant language impairment plays a role in at least some speech perception deficits in children with CAS, which is reminiscent of findings that concurrent receptive language impairment in children with dyslexia is a quasi-measure of their speech perception difficulty (Robertson et al., 2009).

Higher-order speech perception studies of children with CAS have also had mixed findings. Children with CAS performed more poorly than those with TD in identification of vowels from two continua (Maassen et al., 2003) and in judging word similarity based on vowel differences (Marion et al., 1993). However, children with CAS and those with TD classified words from real word pair continua (e.g., a [bak] 'box' to [dak] 'roof' continuum in Dutch) similarly (Groenen et al., 1996; Nijland, 2009). On a highly metalinguistic task involving rhyming, children who had CAS and receptive language within normal limits nonetheless performed significantly worse than those with TD on determining whether nonword pairs rhymed (Nijland, 2009), generating rhymes for real words, selecting the best rhyme for a given real word, and judging which of a set of 10 words rhymed with a given real word (Marion et al., 1993). Overall, these results are suggestive of speech perception deficits among at least some children with CAS, at least in comparison to those with TD.

There have been few investigations of the perception of suprasegments in children with CAS. The suprasegmental aspects of speech that have been explored include vowel duration and pitch. Ingram et al. (2019) found that children with CAS (without controlling for language ability) were less accurate in discriminating vowel durations than TD children. Nijland (2009) attempted to examine the high and low pitch discrimination skills of TD children and children with CAS, phonologically disordered (PD), and "mixed" (symptoms of both PD and CAS). All of the children were asked to perform a frequency pattern task in which a sequence of three tones

was presented. One of the tones was different from the other two with respect to frequency. The children were to indicate which tone differed. However, the author reported that only 19% of children with SSDs (CAS and/or PD) finished the task even though practice trials were given, compared with an 80% completion rate among the TD children. Nijland (2009) attributed the task completion failures to deficient short-term memory in children with SSDs, which was demonstrated via another task in the same study. Nijland hypothesized that the children were not able to store the sounds in memory to determine which sound was different from the others. The nonsense nature of the high and low frequency stimuli did not provide any semantic information that would facilitate memory, thus resulting in heavy cognitive loads.

The current understanding of speech perception skills in children with CAS is mainly based on English speakers. Similar to other Indo-European languages, English has lexical stress patterns that are formed via simultaneous variations of three physical parameters, namely fundamental frequency (F0), duration, and intensity. These three parameters contribute to the perception of pitch, length, and loudness respectively. When F0 is varied in a lexically stressed syllable, duration and/or intensity usually vary at the same time. Lexical meanings cannot be changed by keeping the segments constant while varying pitch only. Therefore, in the context of speech, only nonsense syllables can be used to test pitch perception skills in English.

Cantonese as a medium for examining pitch perception skills

Fortunately, there is a way to examine speech-related pitch perception skills in children with CAS who speak Cantonese. The problem of the heavy cognitive load of pitch perception tasks (as in Nijland, 2009) can be solved by using words with lexical tones as stimuli. Lexical tones are present in tonal languages, such as Mandarin, Cantonese, Thai and Vietnamese. Lexical tones are formed using variations of fundamental frequency to indicate meaning. In tonal

languages, the same segments can be realized with different F0 patterns in terms of levels and/or contours to form different lexical tones. Cantonese, a tonal language, has six contrastive tones (Yip & Matthews, 2011): Tones 1, 3 and 6 are high-, mid- and low-level tones, respectively. Tone 2 is high rising, tone 5 is low rising and tone 4 is a low falling tone. In English orthography, a number after the syllable indicates the tone of that syllable, as in /jɐu4 sœy2/ ('swim'; Chinese character: 游水), which has a low-falling tone on the first syllable and a high-rising tone on the second. The variations in F0 for the six tones indicate different lexical meanings. For example, /jɐu4 sœy2/ (Chinese character: 游水) means 'swim' and /jɐu4 sœy3/ (Chinese character: 游說) means 'convince.' The two verbs share identical segmental features; the tones differentiate their lexical meanings.

Using words with lexical tones as the stimuli in a pitch perception task is a useful research strategy. Children can be asked to identify specific pitch variations in auditory stimuli by giving a non-verbal response (e.g. picture pointing), mapping the auditory information to meanings that are stored in lexical representations. This mapping reduces the memory demand. Wong et al. (2009) examined tone perception skills in preschool children with and without development language disorders by using contrastive word pairs. Their successful investigation demonstrates that pitch perception skills can be investigated even in young children.

Tone perception and production in Cantonese-speaking children with and without CAS

Previous research has investigated tone perception skills in Cantonese-speaking children using the Cantonese Tone Identification Test (CANTIT; Lee, 2012). This standardized assessment is designed as a picture-pointing task with 15 combinations of tone pair contrasts (e.g., tone 1 vs. tone 2, tone 1 vs. tone 3). Normative data from three-year-old children to adults are provided in seven age groups, with 25 participants in each age group. Given that vocabulary

knowledge has an impact on tone identification skills in Cantonese-speaking preschool children (Wong et al., 2009), the test stimuli were assessed by a group of experts, whose consensus was that they are appropriate for children with receptive vocabulary age equivalents of at least three years (Lee, 2012). Research using this standardized assessment has indicated that tone 1 (high level) and tone 5 (low rising) are the easiest and the most difficult tones for 3-year-old TD children to identify, respectively (Lee et al., 2015; Wong et al., 2017). Another study extended the investigation to TD children aged from four to six years of age (Wong & Leung, 2018). Although 4-year-old children showed the lowest perception accuracy on tone 3 (mid-level), 5- and 6-year-old children found tone 5 (low rising) the most difficult tone to identify (Wong & Leung, 2018). Regarding error patterns, TD 3- to 6-year-old children confused the high-rising tone 2 with the low-rising tone 5, the mid-level tone 3 with the low-level tone 6 (Lee et al., 2015; Wong et al., 2017; Wong & Leung, 2018), the low-falling tone 4 with the low-level tone 6 (Lee et al., 2015) and the high-rising tone 2 with the low-falling tone 4 (Wong et al., 2017; Wong & Leung, 2018).

With respect to tone production skills, Wong et al. (2017) suggested that, when lexical priming effects on the listeners are controlled, children at age three do not produce the six Cantonese tones with adult-like accuracy. The descending order of production accuracy of the six tones was tone 5 (low rising), tone 4 (low falling), tone 1 (high level), tone 2 (high rising), tone 3 (mid level) and tone 6 (low level), with the first three easier than the last three tones. Wong and Leung (2018) extended the findings to children from ages four to six years. They found that no 4-year-old children produced any of the tones in an adult-like manner. Five-year-old children produced two tones (i.e., tone 5 and 6) and 6-year-old children produced three tones (i.e., tone 4, 5 and 6) with adultlike accuracy. These studies demonstrated that tone production

skills emerge in children starting at 5 years of age. However, in real clinical setting, tone production skills are usually tested using a standardized test, the Hong Kong Cantonese Articulation Test (HKCAT, Cheung et al., 2006), which includes a set of normative data for children (aged from 2;4 to 11;7) and adults (aged between 18 to 45) with about 120 participants in each group. During the test, clinicians are required to perceptually judge the children's productions by comparing them with the expected productions. Potential lexical priming effects on listeners are therefore involved in this procedure. In this situation, children are perceived to produce all Cantonese lexical tones before age 2;6 (Cheung et al., 2006; So & Dodd, 1995; To et al., 2013; Tse, 1978).

The predictive relationship between tone perception and production accuracy in TD Cantonese-speaking children aged between three to six years old has also been explored, by Wong et al. (2017) and Wong and Leung (2018). No significant correlation was found in 3-year-old children and a weak correlation was reported in 4- to 6-year-old children. It was observed that there were discrepancies between tone perception and production accuracies, for example: a tone that was perceived correctly may not be produced correctly, and the tone with the lowest perception accuracy was not the tone with the lowest production accuracy. These authors suggested that accurate tone perception is not sufficient for accurate tone production. They proposed that other factors such as reduced speech motor control may account for lexical tone production errors (Wong et al., 2017; Wong & Leung, 2018).

There are fewer studies of tone production or perception in Cantonese-speaking children with CAS. Wong et al. (2021) reported a deep investigation of tone production skills in children with CAS in the context of a tone sequencing task which required children to repeat one-, two-, and three-syllable words with one or more Cantonese tones (i.e. tones 1, 2 and 4). They reported

that, compared with children with non-CAS speech and/or language impairment and TD children, children with CAS showed a flatter slope of F0 values on high-rising tones (tone 2) and a higher F0 value with a similar falling pattern for low-falling tones (tone 4). They suggested that children with CAS have more difficulty varying their vocal pitch for tone than children without CAS. As children with CAS have difficulty in producing the high-rising tone (tone 2), they tend to use the high-level tone (tone 1) as a substitute.

There is also a lack of understanding of the relationship between tone perception and tone production in children with CAS. With respect to their tone perception, Wong and Lee (2018) reported the performance of two Cantonese-speaking children with CAS on the CANTIT. Both participants demonstrated Z scores lower than -1.35, which indicated a deficit in tone identification skills. With respect to their errors, no specific pattern was found when all possible tone pair contrasts were considered. Although the children showed poor test results, it is unknown whether their degraded lexical tone identification skills were impacted by their co-occurring poor receptive vocabulary skills, by deficits in F0 processing, or both. To understand pitch perception skills and their relationship to tone production skills in children with CAS, an investigation of participants who have sufficient vocabulary knowledge is needed.

Thus, the purpose of this study was to investigate tone identification skills and the relationship between lexical tone perception and production in children with CAS. Prior assessments of receptive vocabulary skills were conducted to rule out this possible confounding factor. The specific hypotheses developed for this study were:

1. Some Cantonese-speaking children with CAS would have similar tone identification skills as their age-matched TD peers, given that poor speech perception may not be a core deficit of CAS (Zuk et al., 2018).

2. Children with CAS would demonstrate similar patterns of perception accuracy among the six Cantonese tones as TD peers, i.e. having the most difficulty in perceiving tone 3 and tone 5 (Wong & Leung, 2018).
3. Children with CAS would demonstrate similar confusion on perceiving Cantonese lexical tones as TD peers, i.e., showing confusion on tone 2 vs. tone 5, tone 3 vs. tone 6, tone 4 vs. tone 6, and tone 2 vs. tone 4 (Lee et al., 2015; Wong et al., 2017; Wong & Leung, 2018).
4. No significant correlations between lexical tone perception and production would be found in children with CAS as accurate tone perception may not be sufficient for accurate tone production. Other factors such as reduced speech motor control may account for lexical tone production errors (Wong et al., 2017).

Methods

This study was approved by the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee (Reference Number: 2018.265).

Participants

The inclusion criteria for the participants were 1) aged between 3;0 and 6;11, 2) Cantonese as the main language used in daily life, 3) no diagnosis of autism spectrum disorder or other biomedical conditions that affect speech, 4) no structural abnormalities that affect speech, and 5) diagnosed or suspected with CAS by a qualified speech-language therapist (SLT). Children who did not imitate sounds were excluded. Three Cantonese-speaking children with CAS (two boys and one girl) participated in this study. Their ages ranged from 3;7 to 5;8 (years;months; M = 54.33 months, SD = 12.67 months). The same children also participated in another study by Wong et

al. (2021).

Procedure

Initial assessment

All children were assessed by the first author, a Cantonese-speaking SLT who has more than 10 years of experience in the differential diagnosis of paediatric speech sound disorders. The assessment protocol comprised a collection of case history, a hearing screening, a speech and language sample, a standardized receptive vocabulary test (Hong Kong Cantonese Receptive Vocabulary Test [HKCRVT; Lee et al., 2009]), a standardized articulation test (Hong Kong Cantonese Articulation Test [HKCAT; Cheung et al., 2006]) and a 2-hour motor speech assessment battery which comprised an assessment of the ability to imitate polysyllabic words or phrases and an oral motor assessment that included ~~a~~ diadochokinesis (DDK). All children passed a 20-dB pure-tone screening at 1000, 2000, and 4000 Hz (American Speech-Language-Hearing Association [ASHA], n. d.). Table 1 lists their assessment results. All the participants were diagnosed with severe speech sound disorders consistent with CAS (more details are provided in the next section). No symptoms of childhood dysarthria were observed. Two participants (A and C) had adequate receptive vocabulary skills while one participant (B) had a moderate deficit in receptive vocabulary skills. Nevertheless, age-equivalent scores showed that all participants had receptive vocabulary skills (> 3;0) adequate to understand all of the stimuli in the Cantonese Tone Identification Test (CANTIT; Lee, 2012).

(Insert Table 1 here.)

Confirmation of CAS Diagnoses

The first author confirmed the CAS diagnoses of all of the participants based on criteria modified from Murray et al. (2015) for Cantonese. Particularly, the children showed five clinical features across four tasks. The features were 1) inconsistent errors on consonants and vowels in repeated productions of syllables or words, 2) lengthened and disrupted coarticulatory transitions between sounds and syllables, 3) lexical tone errors, 4) reduced accuracy in multisyllabic words, i.e. the percentage of phonemes correct (PPC) in polysyllabic words or sentences is smaller than 60% and 5) reduced phone accuracy on the DDK task, i.e. the PPC in stimuli with three different syllables on DDK tasks is smaller than 70% . The working definitions of the clinical features were presented in Wong et al. (2021) in detail. The four tasks were 1) speech sample, 2) HKCAT, a standardized articulation test for Cantonese speakers, 3) DDK and 4) imitation of polysyllabic words. The first three clinical features were each observed at least once on all of the tasks except inconsistent errors on the speech sample. This is because children may not intend to produce the same stimuli several times in their spontaneous speech, as is needed for the consistency measure. The fourth clinical feature was observed at least once on all of the tasks except the DDK, while the fifth clinical feature was observed on the DDK only.

Data collection

The tone perception skills of the children were tested using the research version of the Cantonese Tone Identification Test (CANTIT-75; Lee, 2012). The test includes 75 monosyllabic words and was presented on a Windows computer. On each trial, the children were required to point to one among four pictures including the target word, a tone distractor (tone minimal pair with the target word), a consonant distractor (same initial consonant as the target word) and a vowel

distractor (same vowel as the target word). The stimuli were recorded by a male speaker in the sentence-final position of the carrier phrase: "邊幅係__ (Which picture is __?)." Performance on each item was recorded automatically by the programme. All the participants completed the test within 20 minutes. Raw scores, Z-scores, percentile scores and error patterns were recorded for later analyses.

The articulation skills of the children were tested in the initial assessment using the HKCAT, which is a commonly used assessment tool used with young children through adults in Hong Kong. The test consists of 50 items presented in 41 photos. The children were required to imitate each item and the assessor (first author) perceptually judged the children's productions as either correct or incorrect, i.e., as matching or not matching the target words. Twelve items, two for each of the six Cantonese tones, were selected for tone production analysis, as these twelve items are used for scoring tone production accuracy on the HKCAT according to the test manual. All of the participants completed the test within 10 minutes. Raw scores, Z-scores and percentile scores were recorded for later analysis.

Data analyses

Children's performance on the perception and production tests were compared with the normative data from the tests. Regarding lexical tone perception, the accuracy of the six tones were compared using a Kruskal-Wallis Test to determine whether the children perceived the six tones differently. Descriptive statistics were also calculated. Based on Wong et al. (2017), 10% difference was set as the criterion for significance of error patterns. The relationship between lexical tone perception and production accuracy was examined using a two-way repeated-measures ANOVA to determine whether children performed similarly on tone perception and production. The within-subjects factor are task (perception or production) and tone (tone 1 to

tone 6). The number of correct items for each tone stimulus was transformed to a percentage for statistical analyses. A Pearson correlation coefficient was used to determine whether tone perception and production were correlated. The first author re-listened the children's productions three months later as an intra-rater reliability measure. An independent SLT who has more than 10 years of experience in Cantonese paediatric speech sound disorders also was invited to judge the lexical tone correctness of the children's productions as an interrater reliability measure.

Results

Intra- and inter-rater reliability

Intra-rater and inter-rater agreement levels on perceptual ratings of Cantonese lexical tones were calculated using Cohen's Kappa coefficient. Both intra- and inter-rater agreement levels were strong (intra-rater: k = .827, $p < 0.001$; inter-rater: k = .827, $p = 0.001$).

Tone perception skills

Participants' tone perception ability, as assessed using the standardized Cantonese Tone Identification Test (CANTIT; Lee, 2012), are presented in Table 2. The test results include how often each child selected the target versus each of the distractors. Child C showed within-normal-limits tone identification skills. Child A showed mild and child B showed moderate difficulties on tone identification skills compared to age-matched TD peers.

(Insert Table 2 here.)

Table 3 shows the tone identification accuracy of each tone stimulus by the participants. All children performed the best on the tone 1 (high-level) stimuli. The two children (children A and B) who had poor tone identification skills performed the worst on tone 3 (mid-level) stimuli,

while the child with age-appropriate tone identification skills (child A) performed the worst on tone 6 (low-level) stimuli. No statistical differences were found among the six tones with respect to perception accuracy, $H(5) = 3.731$, $p = 0.589$.

(Insert Table 3 here.)

Table 4 is a confusion matrix of the tone responses of the three participants. There was one error pattern that occurred over 10% of the time: tone 3 (mid-level) stimuli were misperceived as tone 1 (high-level) stimuli. In addition, the participants misperceived tone 1 (high-level) as tone 6 (low-level) stimuli; tone 2 (high-rising) as tone 4 (low-falling) stimuli; tone 4 (low-falling) as tone 3 (mid-level), 5 (low-rising) and 6 (low-level) stimuli; tone 5 (low-rising) as tone 6 (low-level) stimuli; and tone 6 (low-level) as tone 2 (high-rising) stimuli.

(Insert Table 4 here.)

Tone production skills

Table 5 displays the tone production accuracy of each selected HKCAT tone stimulus by the participants. The children produced tone 1 (high-level) and tone 3 (mid-level) stimuli the best, while tone 5 (low rising) stimuli were the most challenging among the six Cantonese tones. The descending order of production accuracy was tones 1 and 3, tones 2 and 6, tone 4 and tone 5.

(Insert Table 5 here.)

Relationship between tone perception and production skills

A two-way repeated-measures ANOVA was conducted to determine if the children performed similarly on tone perception and production. There were no main effects for task ($F(1,2) = 0.040$, $p = 0.859$) or tone ($F(5,10) = 0.997$, $p = 0.467$); nor were there task or tone interaction effects on perception versus production test accuracy ($F(5,10) = 1.772$, $p = 0.206$). A Pearson correlation

coefficient determined that tone perception and production accuracy were not significantly correlated ($r^2 = 0.181$, $p = 0.078$).

Discussion

The purpose of this study was to investigate tone identification skills and the relationship between lexical tone perception and production in children with CAS, with four proposed hypotheses.

Tone perception skills

The first hypothesis, that some Cantonese-speaking children with CAS would have similar tone identification skills as their TD peers, was supported. The results of the CANTIT showed that two out of three participants had poorer tone identification skills than their peers, while one participant demonstrated age-appropriate skills (Table 2). Deficits in tone identification are not present in every child with CAS. This suggests that poor tone identification may not be a core deficit in Cantonese-speaking children with CAS, thus supporting the suggestion by Zuk et al. (2018) that poor speech perception is not a core deficit of CAS.

The vocabulary knowledge of the participants with CAS in this study was assessed and considered to be appropriate for understanding all of the stimuli on the tone identification test. Therefore, the difficulty in tone identification found in the participants with CAS may be attributed to factors other than poor vocabulary knowledge. It was noted that, on the CANTIT (Table 2), the participants selected not only tone distractors but also vowel and consonant distractors. This leads to the hypothesis that the participants may have had difficulty in identifying consonants and vowels when tones are involved, given that children with CAS do not

have significant problems in perceiving segments in words compare with their TD peers (Bridgeman & Snowling, 1988; Groenen et al., 1996; Nijland, 2009).

Regarding perception accuracy, there were no statistically significant differences among the six Cantonese tones. The second hypothesis was partially supported. The descending order of perception accuracy by the participants was tone 1, tone 2, tone 5, tone 6, tone 4 and tone 3 (Table 3). The easiest tone for the participants to identify was tone 1, which matches well with previous findings (Lee et al., 2015; Wong et al., 2017; Wong & Leung, 2018). The most difficult tone for the participants to identify was tone 3. This is consistent with Wong and Leung (2018), which indicated that tone 3 was the most difficult tone for 4-year-old TD children. However, tone 5 was relatively easier for the children with CAS in this study than for those with TD in that study. The similarities with respect to the easiest and the most difficult tones to perceive across groups suggest that there may be a similar pitch perception mechanism shared by children with CAS and their TD peers.

However, the error patterns found in this study revealed that the children with CAS had more confusions on perceiving tones than TD children. In previous research, TD children showed confusion on similar tones, i.e., those with similar contours but different levels (such as high-rising tone 2 versus low-rising tone 5 and mid-level tone 3 versus low-level tone 6) or similar initial levels but different contours (such as low-falling tone 4 versus low-level tone 6) (Lee et al., 2015; Wong et al., 2017; Wong & Leung, 2018). The participants in this study shared one frequent confusion with the TD children (mid-level tone 3 versus low-level tone 6; 7.69% occurrence) but with two additional frequent tone perception errors (mid-level tone 3 versus high-level tone 1 with 10.26% occurrence and low-rising tone 5 versus low-level tone 6 with 7.69% occurrence). These tone pairs do share a similar contour and a similar starting onset

frequency, respectively. In addition, other frequent tone perception errors (7.69% - 8.33% occurrence) among the children with CAS were also found on less similar tone pairs, which differed on dimensions of both starting level and contour (low level tone 6 versus high rising tone 2, mid-level tone 3 versus high-rising tone 2, high-rising tone 2 versus low-falling tone 4, and mid-level tone 3 versus low-falling tone 4). Though confusion between high-rising tone 2 and low-falling tone 4 can be found in 3- and 4-year-old TD children (Wong et al., 2017; Wong & Leung, 2018), more non-similar confusions were found in our sample of children with CAS. Thus, the children with CAS may make more perception errors on tone pairs that are less similar to each other as well as on pairs that are similar to each other on one dimension, than their TD peers. The third hypothesis, that children with CAS would demonstrate similar perceptual confusions on Cantonese lexical tones as TD peers, was therefore partially supported. Further investigation is needed to understand the possible factors affecting lexical tone identification skills in children with CAS.

In summary, the preliminary findings on tone perception indicate that, despite the group differences, not all Cantonese-speaking children with CAS have a deficit in tone identification skills; there may be other factors affecting these skills.

Tone production skills

The current study reported the results of lexical tone production in three Cantonese-speaking children with CAS on the HKCAT. Table 5 shows that no one tone had a mean of 100% accuracy. Since the current study involved making perceptual judgements of tone correctness with knowledge of the target words, the lexical priming effects on the listeners were not controlled. Given that Cantonese speakers are expected to produce all Cantonese tones before age 2;6 under such circumstances (Cheung et al., 2006; So & Dodd, 1995; To et al., 2013; Tse,

1978), the results suggest that children with CAS have lexical tone production difficulties. This may be due to their underlying speech motor planning and/or programming impairment. Children with CAS not only have difficulty producing high-rising and low-falling tones (Wong et al., 2021) but also on other Cantonese lexical tones in the context of monosyllabic and multisyllabic words (i.e., the test stimuli on the HKCAT).

Relationship between tone perception and production

This study also investigated the relationship between tone perception and production in children with CAS. No main or interaction effects were found and the correlation between lexical tone perception and production accuracy was not significant. This suggests that children with CAS do not perceive lexical tones better or worse than they produce them. The last hypothesis was supported. Compared with young TD children who have better tone identification than production skills (Wong et al., 2017), the current study suggests that, for children with CAS as well, tone identification and production skills may involve processing mechanisms that are independent in at least some respects, such that accurate tone perception skills do not guarantee accurate productions of lexical tones.

In summary, in this small sample, tone production was more universally affected than tone perception.

Clinical implications

The current study informs clinical practice. First, clinicians who work with children with CAS should pay attention to both tone perception and production skills. Although accurate tone perception skills do not guarantee accurate tone production in children with CAS, degraded tone perception skills should be targeted as they affect communication. Given that varied F0 levels

and contours in tonal languages indicate lexical meaning, incorrect perception or production of lexical tones may result in misunderstandings. Therefore, a comprehensive assessment protocol for tonal language speakers with CAS should include pitch perception or tone identification tests as well as tone production tests. Moreover, given that most of the evidence-based treatment approaches for children with CAS (such as Dynamic Temporal and Tactile Cueing [Strand, 2020] and Rapid Syllable Transition Treatment [McCabe et al., 2020]), require high-frequency practice of segmental and suprasegmental forms, degraded pitch perception skills may affect the course of treatment or even treatment outcomes.

Limitations and future investigations

This study has several limitations. First, due to the small sample size of children with CAS, caution should be used when generalizing the results to the population. In future studies, more participants should be recruited to increase statistical power. Second, the current investigation was based on results from two standardized tests, one for tone identification and one for tone production. However, the numbers of stimuli on the tests are not the same. Testing the same number of items is recommended for further investigations to better balance the perception and production tasks. Third, the current study focused on higher-order tone identification. It is recommended that researchers look at lower-order tone discrimination skills as well to have a wider understanding of pitch perception skills in children with CAS. Notwithstanding these limitations, this study is the first-ever attempt to investigate lexical tone perception and its relationship with tone production in children with CAS.

Acknowledgements

The authors are sincerely grateful to all the participants and their families as well as to Ms.

Christine So for completing the interrater reliability measures.

Declaration of interest statement

There are no conflicts of interest reported by the authors.

Funding

This study was supported by the Office of Research and Knowledge Transfer Service at the Chinese University of Hong Kong and the Social Innovation and Entrepreneurship Development Fund (SIE Fund), which was granted to the first, fourth and last authors. The reference number is KPF18HLF26.

References

American Speech-Language-Hearing Association [ASHA]. (2007). *Childhood apraxia of speech [Technical Report]*. https://www.asha.org/policy/tr2007-00278/

American Speech-Language-Hearing Association [ASHA]. (n. d.). *Childhood hearing screening*. Retrieved 16 August from https://www.asha.org/practice-portal/professional-issues/childhood-hearing-screening/#collapse_1

Berti, L. C., de Assis, M. F., Cremasco, E., & Cardoso, A. C. V. (2021). Speech production and speech perception in children with speech sound disorder. *Clinical Linguistics & Phonetics*, 1-20. https://doi.org/10.1080/02699206.2021.1948609

Bridgeman, E., & Snowling, M. (1988). The perception of phoneme sequence: A comparison of dyspraxic and normal children. *International Journal of Language & Communication Disorders*, *23*(3), 245-252. https://doi.org/https://doi.org/10.3109/13682828809011936

Cabbage, K. L., Hogan, T. P., & Carrell, T. D. (2016). Speech perception differences in children with dyslexia and persistent speech delay. *Speech Communication*, *82*, 14-25. https://doi.org/https://doi.org/10.1016/j.specom.2016.05.002

Cheung, P. S. P., Ng, K. H., & To, C. (2006). *Hong Kong Cantonese Articulation Test*. Language Information Sciences Research Centre and City University of Hong Kong.

Groenen, P., Maassen, B., Crul, T., & Thoonen, G. (1996). The specific relation between perception and production errors for place of articulation in developmental apraxia of speech. *Journal of Speech and Hearing Research*, *39*(3), 468-482. https://doi.org/10.1044/jshr.3903.468

Ingram, S. B., Reed, V. A., & Powell, T. W. (2019). Vowel duration discrimination of children with childhood apraxia of speech: A preliminary study. *American Journal of Speech-Language Pathology*, *28*(2S), 857-874. https://doi.org/doi:10.1044/2019_AJSLP-MSC18-18-0113

Iuzzini-Seigel, J., Hogan, T. P., Guarino, A. J., & Green, J. R. (2015). Reliance on auditory feedback in children with childhood apraxia of speech. *Journal of Communication Disorders*, *54*, 32-42.

Iuzzini-Seigel, J., & Murray, E. (2017). Speech assessment in children with childhood apraxia of speech. *Perspectives of the ASHA Special Interest Groups*, *2*(2), 47-60. https://doi.org/doi:10.1044/persp2.SIG2.47

Lee, K. Y. S. (2012). *Hong Kong Cantonese Tone Identification Test*. Department of Otorhinolaryngology, Head and Neck Surgery. The Chinese University of Hong Kong.

Lee, K. Y. S., 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. https://doi.org/10.3109/17549507.2014.898096

Lee, K. Y. S., Lee, L. W., & Cheung, P. S. P. (2009). *Hong Kong Cantonese Receptive Vocabulary Test*. Institute of Human Communicative Research, The Chinese University of Hong Kong.

Maassen, B., Groenen, P., & Crul, T. (2003). Auditory and phonetic perception of vowels in children with apraxic speech disorders. *Clinical Linguistics & Phonetics*, *17*(6), 447-467. https://doi.org/10.1080/0269920031000070821

McCabe, P., Thomas, D. C., & Murray, E. (2020). Rapid syllable transition treatment: A treatment for childhood apraxia of speech and other pediatric motor speech disorders. *Perspectives of the ASHA Special Interest Groups*, *5*(4), 821-830. https://doi.org/doi:10.1044/2020_PERSP-19-00165

Murray, E., McCabe, P., Heard, R., & Ballard, K. J. (2015). Differential diagnosis of children with suspected childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *58*(1), 43-60. https://doi.org/doi:10.1044/2014_JSLHR-S-12-0358

Nijland, L. (2009). Speech perception in children with speech output disorders. *Clinical Linguistics & Phonetics*, *23*(3), 222-239. https://doi.org/10.1080/02699200802399947

Robertson, E. K., Joanisse, M. F., Desroches, A. S., & Ng, S. (2009). Categorical speech perception deficits distinguish language and reading impairments in children. *Developmental Science*, *12*(5), 753-767. https://doi.org/10.1111/j.1467-7687.2009.00806.x

Shriberg, L. D., Kwiatkowski, J., & Mabie, H. L. (2019). Estimates of the prevalence of motor speech disorders in children with idiopathic speech delay. *Clinical Linguistics & Phonetics*, *33*(8), 679-706. https://doi.org/10.1080/02699206.2019.1595731

So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language*, *22*(3), 473-495. https://doi.org/10.1017/S0305000900009922

Strand, E. A. (2020). Dynamic temporal and tactile cueing: A treatment strategy for childhood apraxia of speech. *American Journal of Speech-Language Pathology*, *29*(1), 30-48. https://doi.org/doi:10.1044/2019_AJSLP-19-0005

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. https://doi.org/doi:10.1044/1092-4388(2012/11-0080)

Tse, J. K.-P. (1978). Tone acquisition in Cantonese: A longitudinal case study. *Journal of Child Language*, *5*(2), 191-204. https://doi.org/10.1017/S0305000900007418

Wong, A. M., Ciocca, V., & Yung, S. (2009). The perception of lexical tone contrasts in Cantonese children with and without specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *52*(6), 1493-1509. https://doi.org/10.1044/1092-4388(2009/08-0170)

Wong, E. C. H., & Lee, K. Y. S. (2018). *Lexical tone identification in childhood apraxia of speech: Two case studies.* Hong Kong Speech and Hearing Symposium 2018, Hong Kong.

Wong, E. C. H., Lee, K. Y. S., & Tong, M. C. F. (2020). The applicability of the clinical features of English childhood apraxia of speech to Cantonese: A modified Delphi survey. *American Journal of Speech-Language Pathology*, *29*(2), 652-663. https://doi.org/10.1044/2019_AJSLP-19-00118

Wong, E. C. H., Velleman, S. L., Tong, M. C. F., & Lee, K. Y. S. (2021). Pitch variation in children with childhood apraxia of speech: Preliminary findings. *American Journal of Speech-Language Pathology*, *30*(3S), 1511-1524. https://doi.org/doi:10.1044/2021_AJSLP-20-00150

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*, *8*, 1-15. https://doi.org/10.3389/fpsyg.2017.01450

Wong, P., & Leung, C. T.-T. (2018). Suprasegmental features are not acquired early: Perception and production of monosyllabic Cantonese lexical tones in 4- to 6-year-old preschool

children. *Journal of Speech, Language, and Hearing Research*, *61*(5), 1070-1085. https://doi.org/doi:10.1044/2018_JSLHR-S-17-0288

Yip, V., & Matthews, S. (2011). *Cantonese: A comprehensive grammar (2nd ed.)*. Routledge.

Zuk, J., Iuzzini-Seigel, J., Cabbage, K., Green, J. R., & Hogan, T. P. (2018). Poor speech perception is not a core deficit of childhood apraxia of speech: Preliminary findings. *Journal of Speech, Language, and Hearing Research*, *61*(3), 583-592. https://doi.org/10.1044/2017_JSLHR-S-16-0106

511 **Tables**

Table 1

Participant assessment results

			HKCRVT			HKCAT				
						IC	V/Di	FC	T	
#	Age	Sex	SD	Age equivalent	Severity	Percentile				Severity
A	3;7	M	-0.5 < SD < 0.0	3;6	WNL	0.1	0.1	0.1	0.1	Severe
B	4;4	M	-2.5 < SD < -2.0	3;8	Moderate	0.1	0.1	0.1	0.1	Severe
C	5;8	F	0.0 < SD < +0.5	5;9-5;10	WNL	0.1	0.1	9	0.1	Severe

Abbreviations. HKCRVT = Hong Kong Cantonese Receptive Vocabulary Test (Lee et al., 2009); HKCAT = Hong Kong Cantonese Articulation Test (Cheung et al., 2006); WNL = within normal limits; IC = initial consonants; V/Di = vowels or diphthongs; FC = final consonants; T = tones; SD = standard deviation.

Table 2

Target and error responses on the Cantonese Tone Identification Test (CANTIT)

Child	Target answer (Raw score)	Tone distractor	Vowel distractor	Consonant distractor	Z score	Percentile	Severity
A	36	15	12	12	<-1.35	<8	Mild
B	38	29	5	3	<-2.24	<4	Moderate
C	66	7	1	1	0	52	WNL

Abbreviation. WNL = within normal limits; CANTIT = Cantonese Tone Identification Test

Table 3

Participants' tone identification accuracy on each tone on the Cantonese Tone Identification Test (CANTIT)

	Tone 1		Tone 2		Tone 3		Tone 4		Tone 5		Tone 6	
Child	(n = 12)	%	(n = 13)	%	(n = 13)	%	(n = 12)	%	(n = 15)	%	(n = 12)	%
A	8	66.67	8	61.54	4	30.77	4	33.33	8	61.54	4	33.33
B	10	83.33	5	38.46	4	30.77	4	33.33	6	46.15	9	75.00
C	12	100.00	13	100.00	11	84.62	11	91.67	11	84.62	8	66.67
Mean	10.00	83.33	8.67	66.67	6.33	48.72	6.33	52.78	8.33	64.10	7.00	58.33
SD	2.00	16.67	4.04	31.09	4.04	31.09	4.04	33.68	2.52	19.36	2.65	22.05

Abbreviation. SD = standard deviation; CANTIT = Cantonese Tone Identification Test.

Table 4

Confusion matrix of tone responses on the Cantonese Tone Identification Test (CANTIT) perception test

	Responses (%)					
Targets	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6
Tone 1	83.33*	2.78	2.78	0.00	0.00	5.56
Tone 2	2.56	66.67*	0.00	7.69	5.13	2.56
Tone 3	10.26	7.69	48.72*	7.69	2.56	7.69
Tone 4	2.78	2.78	5.56	52.78*	5.56	5.56
Tone 5	2.56	5.13	0.00	5.13	64.10*	7.69
Tone 6	2.78	8.33	5.56	5.56	5.56	58.33*

* indicates correct answers.

524

Table 5

Participants' tone production accuracy on tone stimuli on the Hong Kong Cantonese Articulation Test (HKCAT)

	Tone 1		Tone 2		Tone 3		Tone 4		Tone 5		Tone 6	
Child	(n = 2)	%	(n = 2)	%	(n = 2)	%	(n = 2)	%	(n = 2)	%	(n = 2)	%
A	1	50	2	100	1	50	1	50	1	50	0	0
B	2	100	1	50	2	100	0	0	0	0	2	100
C	2	100	1	50	2	100	2	100	0	0	2	100
mean	1.67	83.33	1.33	66.67	1.67	83.33	1.00	50.00	0.33	16.67	1.33	66.67
SD	0.58	28.87	0.58	28.87	0.58	28.87	0.58	50.00	0.58	28.87	0.58	57.74

Abbreviation. SD = standard deviation.