ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Acoustic Analyses of Tone Productions in Sequencing Contexts Among Cantonese-Speaking Preschool Children With and Without Childhood Apraxia of Speech

Eddy C. H. Wong,[a] Min Ney Wong,[a,b,c,d] and Shelley L. Velleman[e]

[a] Department of Chinese and Bilingual Studies, Faculty of Humanities, The Hong Kong Polytechnic University, Hung Hom [b] Research Centre for Language, Cognition, and Neuroscience, The Hong Kong Polytechnic University, Hung Hom [c] Research Institute for Smart Ageing, The Hong Kong Polytechnic University, Hung Hom [d] The HK PolyU-PKU Research Centre on Chinese Linguistics, The Hong Kong Polytechnic University, Hung Hom [e] Department of Communication Sciences and Disorders, The University of Vermont, Burlington

ARTICLE INFO

Article History:
Received June 21, 2023
Revision received November 5, 2023
Accepted February 16, 2024

Editor-in-Chief: Cara E. Stepp
Editor: Nancy Pearl Solomon

https://doi.org/10.1044/2024_JSLHR-23-00383

ABSTRACT

Purpose: Pitch variations (tone productions) have been reported as a measure to differentiate Cantonese-speaking children with and without childhood apraxia of speech (CAS). This study aims to examine fundamental frequency (*F*0) changes within syllables and the effects of syllable structure, lexical status, and syllable positions on *F*0 in Cantonese-speaking preschool children with and without CAS.

Method: Six children with CAS, six children with non-CAS speech sound disorder plus language disorder (S&LD), 22 children with speech sound disorder only (SSD), and 63 children with typical speech-language development (TD) performed the tone sequencing task (TST). Growth curve analysis was employed to analyze and compare the *F*0 values within syllables with three Cantonese tones (high level, high rising, and low falling). The analysis considered the effects of syllable structure (vowel and consonant–vowel), lexical status (word and nonword), and syllable position (initial, medial, and final) on *F*0, as well as comparisons within and between groups.

Results: Within each group, the effects of syllable structure and position on *F*0 values were found with different patterns. Between-group comparisons showed that the CAS group had reduced *F*0 contrasts. The CAS group could be differentiated from the control groups based on interactions of *F*0 with syllable structure and position, but not lexical status. The dissimilarity of *F*0 values detected between the CAS and SSD/TD groups was more prominent than that observed between the CAS and S&LD groups.

Conclusions: This study demonstrated that Cantonese-speaking children with CAS had difficulty in varying *F*0 within syllables as compared to those without CAS, suggesting pitch variation difficulty and language-specific impairment profiles in CAS. Future investigations of objective measures for identifying Cantonese speakers with CAS and cross-linguistic investigations using growth curve analysis and the TST are suggested.

Childhood apraxia of speech (CAS), a type of motor speech disorder with onset in childhood, is characterized by impairment of motor planning and/or programming of speech movements (Shriberg et al., 2019). Three consensus features, inconsistency across repeated productions of the same word (token-to-token inconsistent errors), lengthened and disrupted co-articulatory transitions, and dysprosody, have been proposed as characteristics of CAS in English speakers (American Speech-Language-Hearing Association [ASHA], 2007). Other proposed features include more variable articulatory movements (Grigos et al., 2015), intrusive schwa (Shriberg et al., 2011), syllable segregation,

Correspondence to Eddy C. H. Wong: eddychwong@gmail.com. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

reduced percentage of phonemes correct on polysyllables, and reduced articulatory accuracy on repetition of [pətəkə] (Murray et al., 2015).

The understanding of CAS in children who speak languages other than English, such as Cantonese (Wong, Wong, & Velleman, 2023a; Wong, Wong, Velleman, Tong, & Lee, 2023), is emerging. It appears that clinical features of CAS manifest differently among languages, at least between English and Cantonese (Wong et al., 2020). These differences are likely due to key characteristics of Cantonese phonology that contrast with the phonological features of English. Cantonese has six distinctive tones (Yip & Matthews, 2011). The six tones can be categorized as level (i.e., high level [T1], mid level [T3], and low level [T6]) and contour (i.e., high rising [T2], low falling [T4], and low rising [T5]) tones. Words that share identical segmental features and sequences but different tones have different meanings. For example, /kaːi1 siː5/ means "market" and /kaːi3 siː4/ means "at that time." Research has shown that lexical status has a priming effect on tone production (Zhao & Jurafsky, 2009). Tone productions are also affected by preceding and subsequent tone syllables; that is, there are carryover and anticipatory tonal co-articulation effects within words (Li & Chen, 2016). In fact, fundamental frequency (*F*0) contours have also been shown to vary depending on their position within a sentence in Taiwan Mandarin (Fon & Hsu, 2004). However, in the computational model of speech production, tone information is considered to be a part of a word's motor program rather than as a prosodic parameter (Guenther, 2016). The syllable structure of Cantonese varies from simple single consonant (C) or vowel (V) to more complex consonant–vowel–consonant (CVC) structures. Cantonese syllables are all of equal duration; it lacks the lexical stress contrasts that are challenging for children with CAS who speak English and similar languages such as Dutch (Kopera & Grigos, 2020; Nijland et al., 2003). Wong et al. (2020) proposed that, due to these linguistic differences between English and Cantonese, some of the clinical features reported for English speakers with CAS, such as lexical stress errors and intrusive schwa, might not be found in Cantonese speakers with CAS. Conversely, Cantonese speakers with CAS exhibit lexical tone errors (Wong, Wong, Velleman, Tong, & Lee, 2023) that are not found in English speakers with CAS (Wong et al., 2020).

Although researchers have suggested that English and Cantonese speakers with CAS may have different clinical manifestations, it remains problematic that diagnostic and treatment practices for Cantonese speakers with CAS predominantly rely on the English literature (Wong, Wong, & Velleman, 2023b). Assessment tests or tasks have been reported with diagnostic accuracy to differentiate English- or Dutch-speaking children with CAS from those with speech delay, phonological disorders, or childhood dysarthria, such as the Dynamic Evaluation of Motor Speech Skill (Strand & McCauley, 2019), Shriberg et al.'s (2017) pause marker, Iuzzini-Seigel et al.'s (2017) inconsistency measure, Iuzzini-Seigel et al.'s (2022) Profile of Childhood Apraxia of Speech and Dysarthria checklist, and Thoonen et al.'s (1999) maximum performance tasks (Murray et al., 2021). However, due to linguistic differences, these assessment tasks may not hold the same diagnostic accuracy for Cantonese speakers with CAS. For example, Chung et al. (2023) recently evaluated the syllable repetition task (SRT; Shriberg et al., 2009) for the assessment of Cantonese speakers with CAS versus speech delay or typical speech-language development (TD). They suggested that the SRT might not be an effective tool to differentiate CAS from speech delay or TD in Cantonese-speaking children. This finding supports the need for more language-specific assessment tools. Although some suggestions have been made for CAS assessment protocols in Cantonese (Wong et al., 2021; Wong, Wong, & Velleman, 2023a), no norm-referenced tools are available yet.

Gaining a better understanding of speech production in Cantonese speakers with CAS is important for developing language-specific assessment tools. The presence of lexical tone errors in Cantonese speakers with CAS, as identified by Wong, Wong, Velleman, Tong, and Lee (2023), may suggest a language-specific diagnostic feature. By further investigating lexical tone errors, researchers can potentially identify specific patterns or characteristics that can guide the development of assessment tools tailored to the unique linguistic demands of Cantonese. Lexical tone errors signify a difficulty in pitch variation, which refers to "one's ability to vary *F*0 within or between syllables when speaking" (Wong et al., 2021, p. 1511). Wong et al. (2021) evaluated the perceptual pitch variation skills and acoustic *F*0 values of Cantonese-speaking children with CAS using the tone sequencing task (TST), which requires rapid repetitions of mono- and tritone syllables with different syllable structures and lexical statuses. The results showed that children with CAS produced lower perceptual tone accuracy and consistency when compared with those without CAS (i.e., children with speech and language disorders or TD). There were no within-group or between-group differences in perceptual tone accuracy on syllables with V versus CV structures or on items with versus without lexical meaning. Wong et al. (2021) further used acoustic measures to find that Cantonese speakers with CAS had difficulty in changing *F*0 within syllables. Based on between-group comparisons of *F*0 values at five time points within syllables (0%, 25%, 50%, 75%, and 100%), the authors concluded that children with CAS produced T2 (high-rising) syllables differently compared with those without CAS. In particular, their T2 syllables were similar to their T1 (high-level) syllables in terms of pitch levels and contours.

The results of Wong et al.'s (2021) study were limited as only the *F*0 value (i.e., pitch height) at each time point within the syllable was analyzed and compared. The rate of *F*0 change (pitch slope) and the acceleration or deceleration of *F*0 change within the syllable (pitch curvature) were not studied. In addition to pitch heights, pitch slopes and curvatures are significant acoustic measures as they may influence the perceptual judgment of tone accuracy (Wong et al., 2024). Moreover, although the authors attempted to investigate the effects of syllable structure and lexical status on *F*0/pitch variation in their study, the findings were limited as only perceptual tone accuracy was analyzed; acoustic analysis might have revealed more differences. Additionally, regarding the version of the TST that was used, biases might have been present because of its inclusion of extremely high-frequency syllables in Cantonese (e.g., /maː1/ ["mother"] and /paː1/ ["father"]) and the use of same-segment CV syllables in each sequence (e.g., [maː1maː2maː4]). Finally, there was only one tone sequence on all of the trisyllabic stimuli (i.e., T1T2T4). Given the carryover and anticipatory tonal co-articulation effects that occur in Cantonese (Li & Chen, 2016), the results might have been affected by the unvaried tone–syllable positions. Thus, there is a need for investigation of pitch variations with controlled tone–syllable positions in varied tone sequences.

In order to examine whether *F*0/pitch variation skills can differentiate Cantonese speakers with CAS, the methodological limitations present in the study of Wong et al. (2021) should be addressed. Statistical comparisons of acoustic findings can be carried out using growth curve analysis (GCA), which treats time as a continuous variable and uses regression techniques to model and analyze the time course data (Mirman, 2014). The time course data in this study were the *F*0 changes across a syllable. These *F*0 changes form unique patterns, which characterize different pitch levels and/or contours in Cantonese lexical tones. GCA provided statistical comparisons of *F*0 values (pitch height), the rate of change of *F*0 values across the syllable (pitch slope), and the rate of change of the pitch slope across the syllable (pitch curvature).

The study aims to (a) examine the *F*0 variations of Cantonese-speaking preschool children with CAS, compared with those with non-CAS speech sound disorder plus language disorder (S&LD), those with non-CAS speech sound disorder only (SSD), and those with typical speech-language development (TD), and (b) examine the effects of syllable structure, lexical status, and syllable position on *F*0 within and between groups. Our hypotheses are the following:

- Hypothesis 1: The CAS group will show significant differences in *F*0 production (i.e., height, slope, and curvature) of the three Cantonese tones (high level, high rising, and low falling) compared with the control groups, while comparable performance will be observed among the control groups.
- Hypothesis 2: The CAS group will show significant within-group differences in *F*0 production between V versus CV syllables with the three Cantonese tones, while the control groups will show comparable within-group performance on V versus CV syllables with the three Cantonese tones. With respect to between-group comparisons, the CAS group will produce V and CV syllables with significantly different *F*0 compared with the control groups, while comparable *F*0 on V and CV syllables will be observed among the control groups. Given that the speech of all children with CAS is characterized by difficulty in transitioning articulatory gestures from phoneme to phoneme (ASHA, 2007), it is expected that Cantonese speakers with CAS will have more difficulty in producing *F*0/pitch in CV syllables than in isolated vowels.
- Hypothesis 3: All of the groups will show significant within-group differences in *F*0 production between items with different lexical status (word vs. nonword) with the three Cantonese tones. With respect to between-group comparisons, the CAS group will show significantly different *F*0 on both word and nonword items compared to the three control groups, while comparable performance will be observed among the control groups. Given that lexical status has a priming effect on tone production (Zhao & Jurafsky, 2009), it is anticipated that children with and without CAS will show acoustic differences on *F*0 between word and nonword items, although no significant group differences were reported in a previous study based on perceptual judgments of children's tone accuracy (Wong et al., 2021).
- Hypothesis 4: All of the groups will show significant within-group differences in *F*0 production among syllables in different positions (initial vs. medial vs. final) with the three Cantonese tones. With respect to between-group comparisons, the CAS group will show a greater number of significant differences in *F*0 on syllables in all three of the positions compared to the three control groups, while comparable performance will be observed among the control groups. Given that the pitch of a syllable can be influenced by the preceding and subsequent tone syllables (Li & Chen, 2016), Cantonese-speaking children may produce *F*0/pitch differently in different syllable positions. Given that children with CAS generally

have more variable articulatory movements (Grigos et al., 2015), it is anticipated that Cantonese-speaking children with CAS will show more varied *F*0 productions among syllables in different positions, compared with those without CAS.

Method

This study was approved by the PolyU Institutional Review Board (Reference No. HSEARS202103300007).

Participants

A total of 97 Cantonese-speaking children across four different groups participated in this study. Six preschool children with CAS (four boys and two girls; M_{age} = 4.39 years, SD = 1.07) participated. The inclusion criteria were (a) aged between 3;00 (years;months) and 5;11, (b) speaking Cantonese as the dominant language for daily communication, (c) suspected or diagnosed with CAS by a speech-language pathologist (SLP), (d) no concomitant developmental disorders such as autism spectrum disorder, (e) no structural abnormalities that affect speech such as cleft lip and/or palate, (f) no uncorrected vision impairment, and (g) no hearing loss. Potential participants who did not fulfill the inclusion criteria were excluded, in addition to potential participants who (a) had no ability to imitate sounds, (b) had incomplete assessment or data collection sessions due to behavioral issues or data loss, or (c) were unable to repeat mono- and trisyllables 5 times, resulting in more than 35% missing values in their data sets.

Three control groups also participated. There were six children with non-CAS speech sound disorder plus language disorder (S&LD group; five boys and one girl; M_{age} = 4.68 years, SD = 0.78), 22 children with non-CAS speech sound disorders only (SSD group; nine boys and 13 girls; M_{age} = 4.58 years, SD = 0.78), and 63 children with typical speech-language development (TD group; 31 boys and 32 girls; M_{age} = 4.34 years, SD = 0.79), based on the described inclusion criteria except the diagnosis of CAS. Since a majority of the participants in the CAS group had receptive language disorder (five out of six children), the inclusion of the S&LD group provided a control regarding receptive language abilities. Cognitive abilities were not tested, but there were no parental reports of cognitive deficits in either the CAS or control groups. The median ages of the groups were not statistically different, $H(3) = 2.341$, $p = .505$.

Assessment and Diagnosis

All of the participants in the CAS group received a 2.5-hr speech and language assessment by an SLP (first author). The assessment protocol included (a) a 20-dB pure-tone screening at 1000, 2000, and 4000 Hz (American Academy of Audiology, 1997); (b) standardized language assessments using the Hong Kong Test of Preschool Oral Language (TOPOL; Child Assessment Service, 2020) and the Hong Kong Cantonese Receptive Vocabulary Test (HKCRVT; Lee et al., 2009); (c) a standardized articulation test, the Hong Kong Cantonese Articulation Test (HKCAT; Cheung et al., 2006); (d) a polysyllabic imitation task; (e) an adapted version of the Robbins and Klee (1987) oral mechanism assessment protocol (Wong, Wong, Velleman, Lai, & Cheung, 2023); and (f) a conversational speech sample. The assessment sessions were video-recorded. The results of these assessments of the children with CAS are provided in Table 1. All of the CAS participants passed the hearing screening.

The first author confirmed the CAS diagnoses of the participants based on the diagnostic criteria that have been proposed by Wong and colleagues for Cantonese speakers by combining appropriate features from past studies of English speakers and those symptoms that they have identified in their own studies of Cantonese speakers (Wong et al., 2021; Wong, Wong, & Velleman, 2023a; Wong, Wong, Velleman, Tong, & Lee, 2023). Participants were diagnosed with CAS if they demonstrated (a) the three consensus features proposed by ASHA (token-to-token inconsistency, lengthened and disrupted coarticulatory transitions, and dysprosody; ASHA, 2007) and (b) four additional clinical features proposed by Wong et al. (2021; i.e., syllable segregation, lexical tone errors, reduced percentage of phonemes correct on polysyllables, and reduced articulatory accuracy on repetition of [p^haːt^haːk^haː] with Cantonese appropriate aspirated stops and long low back vowels) across four assessment tasks (i.e., conversational speech sample, polysyllable imitation task, standardized articulation test, and diadochokinetic tasks). The presence of clinical features was supported by predefined thresholds proposed by Wong et al. (2021; i.e., phoneme consistency score < 70%, percent tones correct < 95%, percent phonemes correct in polysyllabic words or sentences < 70%, percent phonemes correct in trisyllabic DDK stimuli < 70%), except syllable segregation and lengthened and disrupted coarticulatory transitions, which were perceptually judged.

The participants in the control groups received a speech and language assessment by either the first author or an SLP student under supervision. The assessment protocol included (a) a hearing screening, (b) the standardized language tests, and (c) the standardized articulation test, as described above. The assessment results were reviewed by the first author, who diagnosed the participants based on the standardized test results. Participants who scored below the 16th percentile on the HKCRVT or any subtests of the TOPOL, including Sentence Comprehension,

Table 1. Assessment results of participants with childhood apraxia of speech.

				TOPOL (percentile)				CRVT	HKCAT (percentile)				HKCAT[a]				PI				DDK
ID	Dx	Age	Sex	SeC	SP	StC[b]	VN		IC	V/D	FC	T	PPC (%)	PTC (%)	PPC-TW (%)	PTC-TW (%)	PPC (%)	CS-P (%)	PTC (%)	CS-T (%)	PPC (%)
1	CAS + LD	5;5	M	0.1	0.1	0.4	0.1	< 0.1	0.1	0.1	0.4	0.1	67.4	92.8	53.3	100	57.8	24.7	52.3	57.1	49
2	CAS + LD	4;7	F	0.4	0.1	0.4	1	24.5	0.1	0.1	9	0.1	65.6	66.7	60	75	60	5.2	60.7	14.3	15.6
3	CAS + LD	5;9	F	0.1	0.1	0.1	0.1	< 0.1	0.1	0.1	0.1	0.1	25.6	27.5	6.7	8.3	6.7	19.5	14.7	23.4	11.1
4	CAS + LD	3;7	M	9	1	NA	63	5.1	0.1	0.1	0.1	0.1	31.5	26.1	26.7	16.7	26.7	7.8	23.2	19.5	29.2
5	CAS	4;0	M	95	50	95	16	89.8	0.1	1	0.1	0.1	49.1	62.3	43.3	50	43.3	18.2	51.6	42.9	57.8
6	CAS + LD	3;0	M	37	< 0.1	NA	16	62.9	9	37	25	0.1	74.4	79.7	63.3	66.7	63.3	3.9	48.5	18.2	35.6

Note. The trisyllabic words on the HKCAT are /kʰʊk1 kʰei4 pɛŋ2/ (“cookie”) and /sɵy5 waːt6 tʰɐi1/ (“slide”). Dx = diagnosis; TOPOL = Hong Kong Test of Preschool Oral Language (Child Assessment Service, 2020); CRVT = Hong Kong Cantonese Receptive Vocabulary Task (Lee et al., 2009); HKCAT = Hong Kong Cantonese Articulation Test (Cheung et al., 2006); PI = polysyllable imitation task; DDK = diadochokinetic task; SeC = Sentence Comprehension subtest; SP = Sentence Production subtest; StC = Story Comprehension subtest; VN = Vocabulary Naming subtest; IC = initial consonant; V/D = vowel or diphthong; FC = final consonant; T = tone; PPC = percent phonemes correct; PTC = percent tones correct; PPC-TW = percent phonemes correct–trisyllabic words; PTC-TW = percent tones correct–trisyllabic words; CS-P = consistency score–phoneme; CS-T = consistency score–tone; CAS = childhood apraxia of speech; LD = language disorder; M = male; F = female; NA = not available.

[a]Additional measures that are not included in the test manual. [b]The Story Comprehension subtest (StC) of the TOPOL was administered to children aged between 4;0 and 5;11 (years;months) according to the test manual.

Sentence Production, Story Comprehension, or Vocabulary Naming, were diagnosed with language disorders. Participants who scored below the 16th percentile on any aspects of the HKCAT, including initial consonants, final consonants, vowels/diphthongs, or tones, were diagnosed with speech sound disorders. All of the participants in the S&LD group had both non-CAS speech sound disorders and language disorders; the participants in the SSD group had non-CAS speech sound disorders only; the participants in the TD group had speech and language abilities within normal limits. All of these participants also passed the hearing screening.

The details of the assessment results of the control groups are presented in Appendix A. The results of group comparisons of the percentiles from the speech and language standardized tests are presented in Appendix B.

Procedure

All of the participants underwent a data collection session, which was held either on the same day or a week after the assessment. During this session, the participants were asked to perform three speech tasks: the SRT (Shriberg et al., 2009), the maximum performance tasks (Thoonen et al., 1999), and the TST. For this study, only the TST was analyzed. The time taken to complete all of the tasks was approximately 60–90 min. Multiple breaks were provided to the participants, if necessary, to ensure that they did not experience fatigue during data collection. The order of the tasks was determined through a random drawing just prior to the commencement of data collection for each participant. Out of the 90 participants, the TST was administered as the first task for 20 participants (22.2%), as the second task for 38 participants (42.2%), and as the final task for 32 participants (35.6%).

TST

The TST used in this study was the third version, which was modified from previous versions (Wong et al., 2019, 2021) and comprised 108 items. The initial and second versions of the TST included seven and 56 items, respectively. The seven items in the first version included mono-, di-, and trisyllabic stimuli, which were formed by initial consonant [m] and vowel [ɔː] with three Cantonese tones (T1, T2, and T4). These three tones were chosen as they encompass the upper and lower frequency limits (i.e., the high and low tones) and all three tone contours in Cantonese (i.e., level, rising, and falling). The 56 items in the second version also included mono-, di, and trisyllabic stimuli, which were formed by [m, p, h, j] and [aː, ɔː, uː, ɛː] with T1, T2, and T4. The TST used in this study included early acquired initial consonants [m, j, t, n]; vowels [aː, ɔː, iː, ɛː]; and T1, T2, and T4. Both the consonants and vowels are acquired before 3;0, and the tones are acquired before 2;6 in Cantonese-speaking children (So & Dodd, 1995). Unlike the previous versions, the current version of the TST avoided the use of extremely high-frequency syllables in Cantonese, such as /maː1/ (“mother”) and /paː1/ (“father”). In addition, the current version of the TST included trisyllabic items with different CV syllables (e.g., [mɛː1jɛː2tɔː4]) and different tone sequences, which were absent from previous versions. Disyllabic items were removed from the current version to keep the length of the task feasible. Thus, the current (third) version of the TST consisted of three sets, including (a) V and CV monosyllabic words and nonwords, (b) V and CV trisyllabic nonwords, and (c) other trisyllabic words and nonwords. The details of these items are provided in Appendices C1, C2, and C3.

The first set was designed to address our hypotheses about the effects of syllable structure and of lexical status on *F*0 variation in syllables with V versus CV structures and in word versus nonword items, respectively. It comprised 24 items, which were derived from four V and four CV structures, combined with three tones. Half of the items were words, and the remaining were nonwords. All of the word items, except for two words (i.e., /ɔː4/ [“goose”] and /nɔː4/ [“move”]) were within the top 4% frequency rank of words in Hong Kong Cantonese-speaking children’s speech, with means and *SD*s of 6.2% and 7.7%, respectively (Lai & Winterstein, 2020). These two words were found to be in the top 21% frequency rank of words in Hong Kong Cantonese-speaking adults’ speech (M = 14.6%, SD = 9.2%; Lai & Winterstein, 2020).

The second set was designed to address our hypotheses related to the effects of syllable structure and syllable position on *F*0 variation in syllables with V versus CV structures and in initial versus medial versus final syllable positions, respectively. It included 72 trisyllabic items, which were derived from three V structures ([aː], [ɔː], [ɛː]), three CV structures ([mɛː], [jɛː], [tɔː]), and three Cantonese tones (T1, T2, T4). Each item included all three tones, forming six different tone sequences: (a) T1T2T4, (b) T1T4T2, (c) T2T1T4, (d) T2T4T1, (e) T4T1T2, and (f) T4T2T1. All of the items were nonwords in Cantonese, and the consonants and vowels used in this set were in the top 50% frequency rank among the 19 consonants and eight monophthongs of Cantonese, while the T1, T2, and T4 tones were at the first, fourth, and fifth/sixth frequency rank among Cantonese tones in three Cantonese corpora, respectively (Andrus et al., 2016; Leung & Law, 2001; Luke & Wong, 2015).

The third set addressed the hypotheses related to the effects of lexical status and syllable position on *F*0 variation in word versus nonword items and in initial versus medial versus final syllable positions, respectively. It included 12 trisyllabic items, six words and six nonwords. All of the words or syllables were found in the lexical database of Hong Kong Cantonese, within the top 3.3% frequency rank of words or syllables in children's speech (*M* = 1.1%, *SD* = 1.1%; Lai & Winterstein, 2020), except for one monosyllable (/tsɔː2/ ["seat"]) and one trisyllable (/taː2 piːn1 lou4/ ["hot pot"]). These words were found to be in the top 7.5% frequency rank of words in Cantonese adults' speech (*M* = 0.64%, *SD* = 0.03%; Lai & Winterstein, 2020). The nonwords were formed using the syllables included in the words. Each item included all three tones in varied order. The difference between the second and third sets was syllable structure. Only V and CV structures were used in the second set, but different within-word syllable structures, ranging from V to CVC, were used in the third set.

The three sets of the TST were randomly presented. Within each set, the items were also presented randomly. The participants were instructed to imitate each item once as a baseline, followed by a sequence composed of five repetitions of the item, in response to the verbal instruction "repeat after me as fast as possible." Regardless of phoneme accuracy, the participants were encouraged to produce sequences consisting of five repetitions of the items. A maximum of three attempts were given for each item. All models of the items and the sequences were prerecorded by a Cantonese-speaking man (the first author) and presented to the participants auditorily via a slide presentation platform with audio capability (Microsoft PowerPoint). The data collection sessions were conducted in soundproof booths using a high-quality microphone (Audio-Technica AT2035 or Shure SM48) and acoustic software interface (Steinberg UR22mkll or M-Audio M-Track Plus II). A sampling frequency of 44 kHz and quantization rate of 16 bits/sample was used.

Data Analyses

Acoustic analyses were conducted to obtain *F*0 values using Praat software (Boersma & Weenink, 2022). Only the sequences of five repetitions of each item were analyzed, not the single-practice productions. The vocalic segment of each syllable was manually identified from a wideband spectrogram in Praat with the aid of audio information by the research assistants who completed annotation in Praat. A Praat script, ProsodyPro (Xu, 2013), was used to identify *F*0 values at 11 evenly spaced time points from the beginning to the end of the vocalic segment of each syllable (0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of the syllable). Following the procedure suggested by Ma et al. (2006), no manual corrections were made. The *F*0 onset values (i.e., 0%) were excluded from the analyses to avoid any possible influence from the initial consonants of the syllable on *F*0. The observed *F*0 was normalized using the *z*-score method (Rose, 1987; Zhang, 2018), which involved calculating the normalized *F*0 values per participant through dividing the difference between the observed *F*0 values and the mean by the standard deviation.

Missing values were observed due to technical issues during data collection, incomplete production of five repetitions, or overlapping productions between the children and assessors. To address randomly missing *F*0 values, predictive mean matching (PMM) imputations were performed. It has been suggested that PMM is a viable method for handling large numbers of numerical missing values (Akmam et al., 2019). This involves estimating the missing *F*0 values based on the specific imputation model. An imputation was conducted for each participant to ensure that the predicted values were calculated from the same participant. The means (and *SD*s) of the percentages of missing values from the CAS, S&LD, SSD, and TD groups were 11.93% (12.49%), 2.50% (4.17%), 2.33% (3.66%), and 3.49% (5.72%), respectively. The imputations were conducted using the Multivariate Imputation by Chained Equations (mice) package (Version 3.14.0).

Effects of Group

GCA (Mirman, 2014) was conducted to examine the time course of normalized *F*0 values across each syllable with one of the three Cantonese tones (T1, T2, and T4) as produced by participants from each of the groups, encompassing all productions from the TST sets. Unlike traditional methods that divide time into distinct intervals, GCA treats time as a continuous variable using regression. It allows for the explicit modeling of the nested structure of the data. The model can include fixed effects, which represent the average patterns of change over time, and random effects, which capture individual differences in the pattern of change. Thus, GCA can be used to estimate parameters such as the level of the variable, the change in the level, the rate of change of the level, and so on (Mirman, 2014).

The overall time course of normalized *F*0 values across a syllable was modeled using a second-order orthogonal polynomial with group (CAS, S&LD, SSD, and TD) serving as a fixed effect at the 10 time points for each tone (10%–100%). Three optimal models were constructed, one for each of the three Cantonese tones, with the CAS group serving as the baseline and parameters estimated for the control groups. The choice of baseline did not restrict the inclusion of the control groups or the comparative analysis between the CAS and control groups as additional pairwise comparisons were conducted. Each model incorporated random participant effects across all time terms. The fixed effects of groups on the intercept, linear term, and quadratic terms were each included separately. The functions are presented as follows:

$$\text{Intercept}: f0 \sim (ot1 + ot2) + Group + (ot1 + ot2|Subject), data = Tone, REML = FALSE \quad (1)$$

$$\text{Linear term}: f0 \sim (ot1 + ot2) + Group + ot1 : Group + (ot1 + ot2|Subject), data = Tone, REML = FALSE \quad (2)$$

$$\text{Quadratic term}: f0 \sim (ot1 + ot2) \times Group + (ot1 + ot2|Subject), data = Tone, REML = FALSE \quad (3)$$

Model comparisons with different terms were conducted to evaluate the effects on model fit, and parameter-specific *p* values were estimated using the normal approximation, treating the *t* values as a *z* value. All analyses were performed using a freeware statistical analysis program (R, Version 4.2.20) with the lme4 package (Version 1.1-29). Pairwise comparisons were conducted to further compare among the groups using multcomp package (Version 1.4-20) after creating a contrast matrix. The R codes employed for the analysis of normalized *F*0 values are provided in Appendix D.

Effects of Linguistic Elements

The effects of the three parameters on normalized *F*0 values between and within groups were examined using GCA (Mirman, 2014). The parameters included syllable structure, lexical status, and syllable position. The overall time course of normalized *F*0 values was modeled using a second-order orthogonal polynomial with group (CAS, S&LD, SSD, and TD), linguistic element (V vs. CV; words vs. nonwords; initial vs. medial vs. final syllable position), and tone (T1, T2, and T4) serving as a fixed effect of the parameters on the intercept, linear term, and quadratic terms individually. A model was constructed with the CAST1CV (participants with CAS producing T1 syllables with CV structure) serving as the baseline and parameters estimated for the same or other groups. The choice of baseline did not restrict the inclusion of the other groups/linguistic elements/ tones or the comparative analysis between the CAS and control groups. Model comparisons with different terms were conducted to evaluate the effects of the parameters on the model fit. Additional pairwise comparisons were performed to examine the differences within group (i.e., CAST1CV vs. CAST1V, CAST2CV vs. CAST2V, CAST4CV vs. CAST4V) and between groups on specific linguistic elements and tones (e.g., CAST1CV vs. S&LDT1CV; CAST1CV vs. SSDT1CV; CAST1CV vs. TDT1CV; and CAST2CV vs. S&LDT2CV, CAST2CV vs. SSDT2CV, etc.) using the multcomp package (Version 1.4-20).

Reliability Measures

Annotations of vocalic segments of syllables were performed by four annotators, research assistants, and undergraduate students who were studying speech and language sciences, after receiving a 2-hr training session from a research assistant with extensive experience in the acoustic analysis of Cantonese speech. Before engaging in independent analyses, all of the annotators were required to achieve good interrater reliability with the trainer as indicated by an intraclass correlation coefficient (ICC) of more than .80 on the *F*0 values of an adult who had had a stroke on the TST. The range of ICC obtained from all of the annotators was between .83 and .90, indicating good interrater reliability with the trainer (Koo & Li, 2016). An ICC was used to calculate the intrarater reliability of each annotator, as well as the interrater reliability of the four annotators. Each annotator independently re-annotated the productions of 10% of the assigned participants, resulting in a total of 10 participants' productions being re-annotated by all four annotators. ICC(2, 1) was used to estimate the level of agreement between two annotations as intrarater reliability. The ICC values of the four annotators were .96 (95% CI [.96, .96], $p < .001$), .99 (95% CI [.99, .99], $p < .001$), 1.0 (95% CI [1.00, 1.00], $p < .001$), and .81 (95% CI [.81, .81], $p < .001$), indicating good to excellent intrarater reliability (Koo & Li, 2016). The interrater reliability among the four annotators was estimated using the ICC(2, 1) as a measure of agreement, and the result indicated good interrater reliability with a score of .86 (95% CI [.86, .86], $p < .001$).

Results

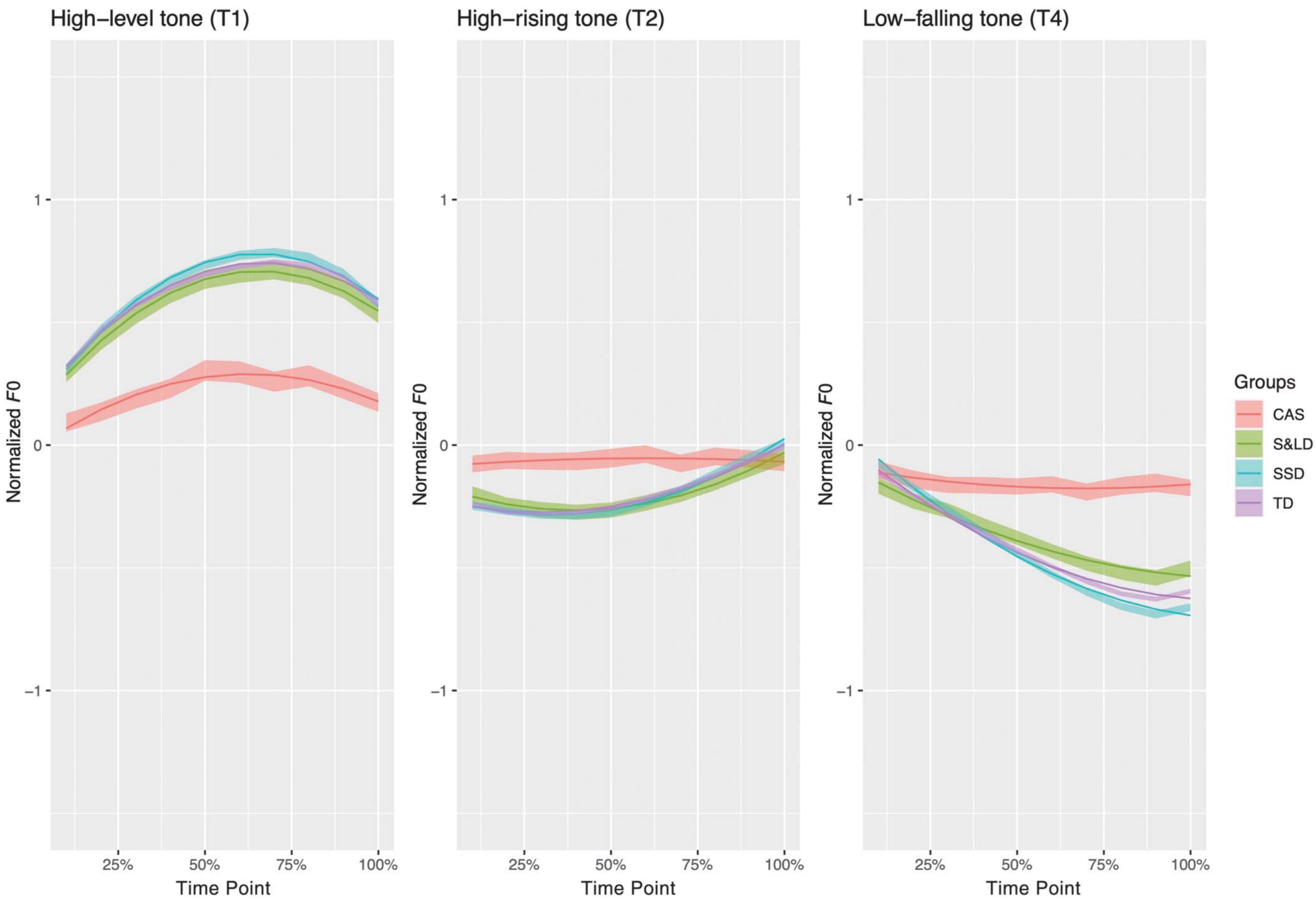
Effects of Group

The model fits for normalized *F*0 values by groups are shown in Figure 1. Table 2 shows the fixed-effect parameter estimates and their standard errors, along with the corresponding *z* values in pairwise comparisons.

High-Level Tone (T1)

For the T1 syllables, the effect of group on the intercept did improve model fits, $\chi^2(3) = 14.53$, $p < .05$, but not that on the linear, $\chi^2(3) = 1.20$, $p > .05$, and quadratic, $\chi^2(3) = 6.98$, $p > .05$, terms, indicating that the groups were different on pitch height, but not pitch slopes nor curvatures. The CAS group produced pitch heights that were statistically significantly lower than those of the SSD and TD groups, but not than those of the S&LD

Figure 1. Growth curve model fits (lines) and observed data (ribbon) for the effect of group on normalized fundamental frequency (*F*0) values on the high-level (T1), high-rising (T2), and low-falling (T4) tones for children with childhood apraxia of speech (CAS), children with non-CAS speech sound disorder plus language disorder (S&LD), children with non-CAS speech sound disorder only (SSD), and children with typical speech-language development (TD).



group, while the control groups had pitch heights that were comparable to each other.

High-Rising Tone (T2)

For the T2 syllables, the effect of group on the linear, $\chi^2(3) = 7.22$, $p < .05$, and quadratic, $\chi^2(3) = 15.82$, $p < .05$, terms did improve model fits, but not that on the intercept, $\chi^2(3) = 1.84$, $p > .05$, indicating that the groups were different with respect to pitch slopes and curvatures, but not pitch heights. The CAS group produced pitch slopes that were statistically significantly shallower than those of the SSD and TD groups, but not than those of the S&LD group. The CAS group exhibited a statistically significant reduction in concave curvature compared to the S&LD, SSD, and TD groups. The control groups had comparable pitch slopes and curvatures.

Low-Falling Tone (T4)

For the T4 syllable, the effect of group on the linear, $\chi^2(3) = 19.94$, $p < .05$, and quadratic, $\chi^2(3) = 9.53$, $p < .05$, terms did improve the model fits, but not that on the intercept, $\chi^2(3) = 2.82$, $p > .05$, indicating that the groups differed on pitch slopes and curvatures, but not pitch heights. The CAS group produced pitch slopes that were statistically significantly shallower compared to the SSD and TD groups, but not to the S&LD group. The control groups had comparable pitch slopes and curvatures.

Effects of Syllable Structure

The model fits for normalized *F*0 values by syllable structure are shown in Figure 2. Table 3 shows the fixed-effect parameter estimates and their standard errors along with the corresponding *z* values in pairwise comparisons. Only the relevant comparisons are presented. The effect of syllable structure on the intercept, $\chi^2(3) = 252{,}646.00$, $p < .05$, and linear, $\chi^2(3) = 20{,}523.80$, $p < .05$, and quadratic, $\chi^2(3) = 8{,}954.10$, $p < .05$, terms did improve model fits, indicating that the *F*0 values of V versus CV syllables were different with respect to pitch heights, slopes, and curvatures.

Table 2. Parameter estimates for analysis of effect of group on normalized fundamental frequency values.

Variable	Pitch height			Pitch slope			Pitch curvature		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
Tone 1 (high-level) syllable									
CAS vs. S&LD	0.179	0.074	2.411	NA	NA	NA	NA	NA	NA
CAS vs. SSD	0.188	0.059	3.171*	NA	NA	NA	NA	NA	NA
CAS vs. TD	0.220	0.055	4.011*	NA	NA	NA	NA	NA	NA
S&LD vs. SSD	0.009	0.059	0.148	NA	NA	NA	NA	NA	NA
S&LD vs. TD	0.041	0.055	0.753	NA	NA	NA	NA	NA	NA
SSD vs. TD	0.033	0.032	1.024	NA	NA	NA	NA	NA	NA
Tone 2 (high-rising) syllable									
CAS vs. S&LD	NA	NA	NA	0.172	0.100	1.723	0.169	0.058	2.923*
CAS vs. SSD	NA	NA	NA	0.268	0.080	3.369*	0.179	0.046	3.882*
CAS vs. TD	NA	NA	NA	0.249	0.074	3.365*	0.151	0.043	3.531*
S&LD vs. SSD	NA	NA	NA	0.096	0.080	1.210	0.010	0.046	0.218
S&LD vs. TD	NA	NA	NA	0.077	0.074	1.037	–0.018	0.043	–0.419
SSD vs. TD	NA	NA	NA	–0.020	0.043	–0.461	–0.028	0.025	–1.128
Tone 4 (low-falling) syllable									
CAS vs. S&LD	NA	NA	NA	–0.338	0.137	–2.471	0.037	0.058	0.646
CAS vs. SSD	NA	NA	NA	–0.596	0.109	–5.459*	0.090	0.046	1.949
CAS vs. TD	NA	NA	NA	–0.481	0.101	–4.754*	0.078	0.043	1.836
S&LD vs. SSD	NA	NA	NA	–0.258	0.109	–2.362	0.053	0.046	1.139
S&LD vs. TD	NA	NA	NA	–0.143	0.101	–1.415	0.041	0.043	0.963
SSD vs. TD	NA	NA	NA	0.115	0.059	1.952	–0.011	0.025	–0.459

Note. *SE* = standard error; CAS = children with childhood apraxia of speech; S&LD = children with non-CAS speech disorder plus language disorder; NA = not available; SSD = children with non-CAS speech sound disorder; TD = children with typical speech-language development.
*indicates $p < .05$.

Within-Group Comparisons

Regarding the within-group comparisons on V versus CV syllables on all three of the Cantonese tones, five and seven statistically significant differences were found in the CAS and each of the control groups (i.e., S&LD, SSD, and TD), respectively, on pitch heights, slopes, and curvatures.

High-level tone (T1). For T1 syllables, the CAS group produced pitch heights and curvatures that were statistically significantly higher and less concave, respectively, on the CV syllables compared to the V syllables, but pitch slopes did not differ based on syllable structure. All three of the control groups produced pitch curvatures that were statistically significantly less convex on the CV syllables compared to the V syllables, but they did not differ with respect to pitch heights or slopes.

High-rising tone (T2). For T2 syllables, the CAS group produced pitch heights that were statistically significantly higher on the CV syllables compared to the V syllables, but pitch slopes and curvatures did not differ. All of the control groups produced pitch heights, slopes, and curvatures that were statistically significantly lower, shallower, and more concave, respectively, on the CV syllables compared to the V syllables.

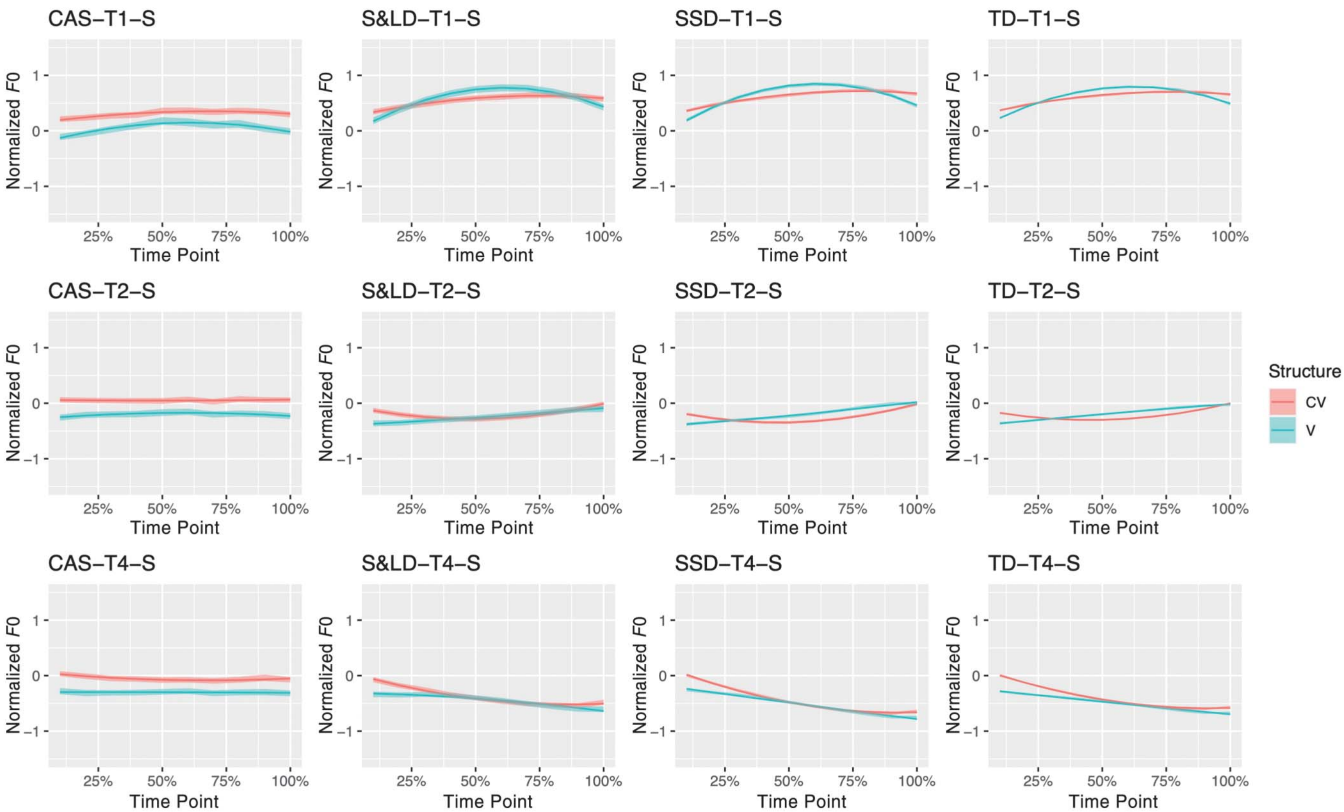
Low-falling tone (T4). For T4 syllables, the CAS group produced statistically significantly higher pitch heights and shallower pitch slopes on the CV syllables compared to the V syllables, but pitch curvatures did not differ. All of the control groups produced pitch heights, slopes, and curvatures that were statistically significantly higher, steeper, and more concave on the CV syllables compared to the V syllables.

Between-Group Comparisons

Regarding the between-group comparisons on V and CV syllables with three Cantonese tones, more statistically significant differences were found between the CAS and each of the control groups, as compared to the comparisons among the control groups with respect to pitch heights, slopes, and curvatures.

High-level tone (T1). For T1 syllables, the CAS group produced pitch heights that were statistically significantly lower on both V and CV syllables compared to all of the control groups; pitch slopes that were statistically significantly shallower on CV syllables compared to the SSD and TD groups, but not to the S&LD group; pitch slopes that were comparable on V syllables compared to all of the control groups; and pitch curvatures that were statistically significantly less convex on V syllables

Figure 2. Growth curve model fits (lines) and observed data (ribbon) for the effect of syllable structure on normalized fundamental frequency (*F*0) values on the high-level (T1), high-rising (T2), and low-falling (T4) tones for children with childhood apraxia of speech (CAS), children with non-CAS speech sound disorder plus language disorder (S&LD), children with non-CAS speech sound disorder only (SSD), and children with typical speech-language development (TD). CV = consonant–vowel; V = vowel.



compared to all of the control groups but were similar on CV syllables. The SSD group produced statistically significant increases in convex curvature compared to the TD group. No statistically significant differences were found for any of the other relevant comparisons between the groups.

High-rising tone (T2). For T2 syllables, the CAS group produced pitch heights that were statistically significantly higher on CV syllables compared to all of the control groups, but similar on V syllables; pitch slopes that were statistically significantly shallower on CV syllables compared to the SSD and TD groups, but not to the S&LD group; pitch slopes that were statistically significantly shallower on V syllables compared to all of the control groups; a statistically significant reduction in concave curvature on CV syllables compared to the SSD and TD groups, but not to the S&LD group; and pitch curvatures that were comparable on V syllables compared to all of the control groups. No statistically significant differences were found for any of the other relevant comparisons between the groups.

Low-falling tone (T4). For T4 syllables, the CAS group produced pitch heights and slopes that were statistically significantly higher and shallower, respectively, on both V and CV syllables compared to all of the control groups, but pitch curvatures did not differ by group. The S&LD group produced pitch slopes (but not pitch heights or curvatures) that were statistically significantly shallower on both V and CV syllables compared to the SSD group, but not to the TD group. The SSD group produced pitch heights and slopes (but not pitch curvatures) that were statistically significantly lower and steeper, respectively, on CV syllables compared to the TD group, but not on V syllables. No statistically significant differences were found for any of the other relevant comparisons among the groups.

Effects of Lexical Status

The model fits for normalized *F*0 values by lexical status are shown in Figure 3. Table 4 shows the fixed-effect parameter estimates and their standard errors along with the corresponding *z* values in pairwise comparisons. Only the relevant comparisons are presented.

Table 3. Parameter estimates for analysis of effect of syllable structure on normalized fundamental frequency values.

Variable	Pitch height			Pitch slope			Pitch curvatures		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
Within-group comparison									
CAST1CV vs. CAST1V	–0.251	0.011	–21.943*	–0.003	0.036	–0.085	–0.143	0.036	–3.951*
CAST2CV vs. CAST2V	–0.253	0.011	–22.145*	–0.106	0.036	–2.941	–0.092	0.036	–2.547
CAST4CV vs. CAST4V	–0.249	0.011	–21.779*	–0.192	0.036	–5.304*	–0.081	0.036	–2.239
S&LDT1CV vs. S&LDT1V	0.032	0.011	2.767	0.011	0.036	0.292	–0.363	0.036	–10.045*
S&LDT2CV vs. S&LDT2V	–0.045	0.011	–3.925*	0.162	0.036	4.477*	–0.218	0.036	–6.020*
S&LDT4CV vs. S&LDT4V	–0.072	0.011	–6.285*	0.131	0.036	3.629*	–0.241	0.036	–6.655*
SSDT1CV vs. SSDT1V	0.017	0.006	2.786	–0.040	0.019	–2.119	–0.403	0.019	–21.367*
SSDT2CV vs. SSDT2V	0.051	0.006	8.606*	0.219	0.019	11.626*	–0.238	0.019	–12.629*
SSDT4CV vs. SSDT4V	–0.073	0.006	–12.278*	0.136	0.019	7.198*	–0.220	0.019	–11.646*
TDT1CV vs. TDT1V	0.008	0.004	2.328	–0.026	0.011	–2.320	–0.311	0.011	–27.873*
TDT2CV vs. TDT2V	0.027	0.004	7.581*	0.179	0.011	16.025*	–0.247	0.011	–22.137*
TDT4CV vs. TDT4V	–0.097	0.004	–27.617*	0.175	0.011	15.647*	–0.197	0.011	–17.661*
Between-group comparison									
CAST1CV vs. S&LDT1CV	0.240	0.027	8.989*	0.139	0.059	2.352	–0.058	0.062	–0.941
CAST1CV vs. SSDT1CV	0.305	0.021	14.319*	0.204	0.047	4.329*	–0.074	0.049	–1.505
CAST1CV vs. TDT1CV	0.299	0.020	15.112*	0.178	0.044	4.064*	–0.066	0.046	–1.436
CAST1V vs. S&LDT1V	0.523	0.027	19.563*	0.153	0.059	2.583	–0.278	0.062	–4.510*
CAST1V vs. SSDT1V	0.573	0.021	26.871*	0.167	0.047	3.545	–0.335	0.049	–6.798*
CAST1V vs. TDT1V	0.558	0.020	28.216*	0.155	0.044	3.542	–0.234	0.046	–5.117*
S&LDT1CV vs. SSDT1CV	0.065	0.021	3.051	0.065	0.047	1.380	–0.016	0.049	–0.326
S&LDT1CV vs. TDT1CV	0.059	0.020	2.965	0.039	0.044	0.885	–0.008	0.046	–0.165
S&LDT1V vs. SSDT1V	0.050	0.021	2.347	0.014	0.047	0.307	–0.056	0.049	–1.144
S&LDT1V vs. TDT1V	0.035	0.020	1.780	0.002	0.044	0.052	0.045	0.046	0.978
SSDT1CV vs. TDT1CV	–0.006	0.011	–0.559	–0.026	0.025	–1.039	0.009	0.026	0.323
SSDT1V vs. TDT1V	–0.015	0.011	–1.294	–0.012	0.025	–0.482	0.101	0.026	3.815*
CAST2CV vs. S&LDT2CV	–0.247	0.027	–9.228*	0.114	0.059	1.935	0.219	0.062	3.543
CAST2CV vs. SSDT2CV	–0.296	0.021	–13.875*	0.176	0.047	3.731*	0.248	0.049	5.048*
CAST2CV vs. TDT2CV	–0.262	0.020	–13.249*	0.167	0.044	3.818*	0.217	0.046	4.758*
CAST2V vs. S&LDT2V	–0.038	0.027	–1.431	0.262	0.059	4.436*	0.093	0.062	1.508
CAST2V vs. SSDT2V	0.009	0.021	0.416	0.381	0.047	8.090*	0.102	0.049	2.075
CAST2V vs. TDT2V	0.018	0.020	0.910	0.332	0.044	7.585*	0.062	0.046	1.366
S&LDT2CV vs. SSDT2CV	–0.049	0.021	–2.306	0.061	0.047	1.305	–0.055	0.036	–1.523
S&LDT2CV vs. TDT2CV	–0.015	0.020	–0.779	0.053	0.044	1.203	0.030	0.049	0.606
S&LDT2V vs. SSDT2V	0.047	0.021	2.210	0.119	0.047	2.529	–0.078	0.036	–2.158
S&LDT2V vs. TDT2V	0.056	0.020	2.844	0.070	0.044	1.591	0.009	0.049	0.184
SSDT2CV vs. TDT2CV	0.034	0.011	2.945	–0.009	0.025	–0.352	–0.031	0.026	–1.179
SSDT2V vs. TDT2V	0.009	0.011	0.796	–0.050	0.025	–1.958	–0.040	0.026	–1.501
CAST4CV vs. S&LDT4CV	–0.324	0.027	–12.147*	–0.373	0.059	–6.308*	0.102	0.062	1.656
CAST4CV vs. SSDT4CV	–0.388	0.021	–18.220*	–0.610	0.047	–12.951*	0.151	0.049	3.075
CAST4CV vs. TDT4CV	–0.342	0.020	–17.293*	–0.518	0.044	–11.856*	0.127	0.046	2.785
CAST4V vs. S&LDT4V	–0.147	0.027	–5.516*	–0.307	0.059	–5.192*	–0.057	0.062	–0.930
CAST4V vs. SSDT4V	–0.213	0.021	–9.977*	–0.540	0.047	–11.452*	0.012	0.049	0.253
CAST4V vs. TDT4V	–0.190	0.020	–9.627*	–0.409	0.044	–9.356*	0.011	0.046	0.244
S&LDT4CV vs. SSDT4CV	–0.064	0.021	–2.993	–0.238	0.047	–5.044*	0.049	0.049	0.999

(*table continues*)

Table 3. *(Continued).*

Variable	Pitch height			Pitch slope			Pitch curvatures		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
S&LDT4CV vs. TDT4CV	−0.017	0.020	−0.879	−0.146	0.044	−3.332	0.025	0.046	0.547
S&LDT4V vs. SSDT4V	−0.065	0.021	−3.062	−0.233	0.047	−4.944*	0.070	0.049	1.420
S&LDT4V vs. TDT4V	−0.043	0.020	−2.173	−0.102	0.044	−2.340	0.069	0.046	1.501
SSDT4CV vs. TDT4CV	0.046	0.011	4.051*	0.092	0.025	3.632*	−0.024	0.026	−0.913
SSDT4V vs. TDT4V	0.022	0.011	1.945	0.131	0.025	5.158	−0.001	0.026	−0.051

Note. *SE* = standard error; CAS = children with childhood apraxia of speech; S&LD = children with non-CAS speech sound disorder plus language disorder; SSD = children with non-CAS speech sound disorder; TD = children with typical speech-language development; T1 = high-level tone; T2 = high-rising tone; T4 = low-falling tone; V = vowel structure; CV = consonant–vowel structure.
**p* < .05.

The effect of lexical status on the intercept, $\chi^2(3) = 65{,}098.40$, $p < .05$, and on the linear, $\chi^2(3) = 3{,}721.00$, $p < .05$, and quadratic, $\chi^2(3) = 1{,}005.90$, $p < .05$, terms did improve model fits, indicating that the *F*0 values of word and nonword items were different regarding pitch heights, slopes, and curvatures.

Within-Group Comparisons

Regarding within-group comparisons, one statistically significant difference was found for each of the CAS, S&LD, and TD groups and two statistically significant differences were found for the SSD group. These significances were related to pitch heights.

Figure 3. Growth curve model fits (lines) and observed data (ribbon) for the effect of lexical status on normalized fundamental frequency (*F*0) values on the high-level (T2), high-rising (T2), and low-falling (T4) tones for children with childhood apraxia of speech (CAS), children with non-CAS speech sound disorder plus language disorder (S&LD), children with non-CAS speech sound disorder only (SSD), and children with typical speech-language development (TD). NW = nonword; W = word.

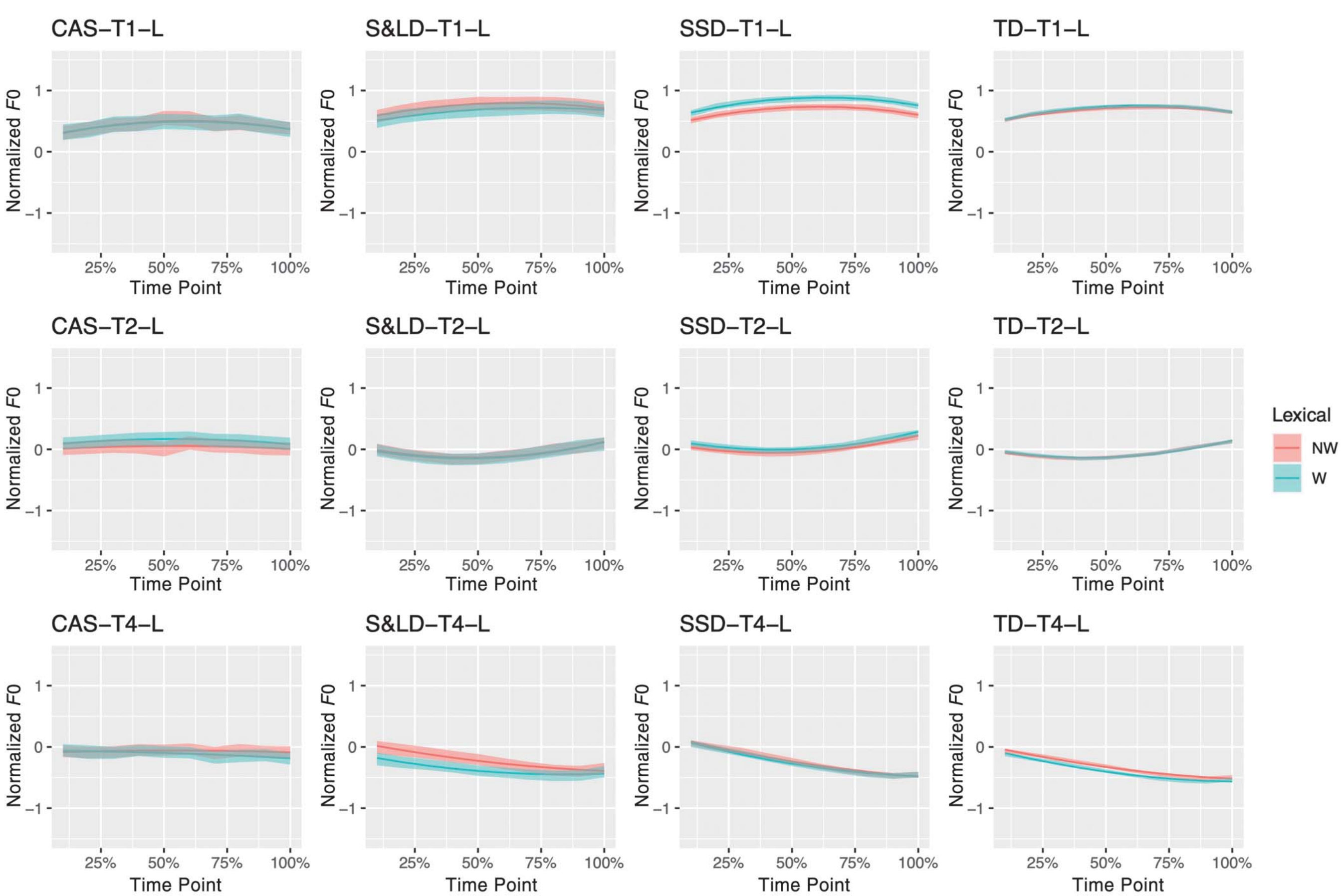


Table 4. Parameter estimates for analysis of effect of lexical status on normalized fundamental frequency values.

Variable	Pitch height			Pitch slope			Pitch curvature		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
Within-group comparison									
CAST1NW vs. CAST1W	–0.007	0.021	–0.337	–0.010	0.068	–0.140	0.023	0.068	0.338
CAST2NW vs. CAST2W	0.099	0.021	4.608*	–0.005	0.068	–0.071	–0.039	0.068	–0.568
CAST4NW vs. CAST4W	–0.042	0.021	–1.962	–0.117	0.068	–1.720	0.007	0.068	0.103
S&LDT1NW vs. S&LDT1W	–0.076	0.021	–3.534	0.056	0.068	0.823	0.043	0.068	0.627
S&LDT2NW vs. S&LDT2W	–0.012	0.021	–0.568	0.014	0.068	0.204	0.000	0.068	–0.004
S&LDT4NW vs. S&LDT4W	–0.142	0.021	–6.605*	0.156	0.068	2.295	0.041	0.068	0.603
SSDT1NW vs. SSDT1W	0.143	0.011	12.801*	0.033	0.035	0.939	–0.011	0.035	–0.298
SSDT2NW vs. SSDT2W	0.056	0.011	4.963*	–0.003	0.035	–0.081	0.017	0.035	0.469
SSDT4NW vs. SSDT4W	–0.023	0.011	–2.055	0.027	0.035	0.754	0.034	0.035	0.951
TDT1NW vs. TDT1W	0.018	0.007	2.716	–0.018	0.021	–0.848	0.013	0.021	0.616
TDT2NW vs. TDT2W	0.002	0.007	0.326	0.011	0.021	0.538	0.029	0.021	1.393
TDT4NW vs. TDT4W	–0.066	0.007	–9.939*	0.062	0.021	2.948	–0.025	0.021	–1.185
Between-group comparison									
CAST1NW vs. SLDT1NW	0.295	0.123	2.400	0.053	0.079	0.664	0.033	0.072	0.453
CAST1NW vs. SSDT1NW	0.224	0.098	2.288	0.023	0.063	0.360	–0.003	0.058	–0.044
CAST1NW vs. TDT1NW	0.229	0.091	2.516	0.064	0.059	1.084	0.041	0.054	0.757
CAST1W vs. SLDT1W	0.226	0.123	1.841	0.118	0.079	1.488	0.052	0.072	0.724
CAST1W vs. SSDT1W	0.375	0.098	3.825*	0.066	0.063	1.036	–0.036	0.058	–0.623
CAST1W vs. TDT1W	0.254	0.091	2.793	0.065	0.059	1.102	–0.007	0.054	–0.134
S&LDT1NW vs. SSDT1NW	–0.071	0.098	–0.720	–0.030	0.063	–0.473	–0.035	0.058	–0.612
S&LDT1NW vs. TDT1NW	–0.066	0.091	–0.727	0.011	0.059	0.186	0.008	0.054	0.144
S&LDT1W vs. SSDT1W	0.149	0.098	1.517	–0.053	0.063	–0.829	–0.088	0.058	–1.531
S&LDT1W vs. TDT1W	0.028	0.091	0.305	–0.053	0.059	–0.909	–0.060	0.054	–1.113
SSDT1NW vs. TDT1NW	0.005	0.053	0.085	0.041	0.034	1.200	0.043	0.031	1.387
SSDT1W vs. TDT1W	–0.121	0.053	–2.295	–0.001	0.034	–0.026	0.029	0.031	0.928
CAST2NW vs. S&LDT2NW	–0.092	0.123	–0.751	0.144	0.079	1.808	0.263	0.072	3.627*
CAST2NW vs. SSDT2NW	–0.012	0.098	–0.120	0.202	0.063	3.183	0.247	0.058	4.281*
CAST2NW vs. TDT2NW	–0.092	0.091	–1.014	0.214	0.059	3.644*	0.240	0.054	4.476*
CAST2W vs. S&LDT2W	–0.203	0.123	–1.655	0.162	0.079	2.043	0.301	0.072	4.155*
CAST2W vs. SSDT2W	–0.055	0.098	–0.561	0.204	0.063	3.214	0.303	0.058	5.236*
CAST2W vs. TDT2W	–0.189	0.091	–2.078	0.201	0.059	3.423	0.291	0.054	5.435*
S&LDT2NW vs. SSDT2NW	0.081	0.098	0.822	0.058	0.063	0.917	–0.015	0.058	–0.265
S&LDT2NW vs. TDT2NW	0.000	0.091	0.000	0.071	0.059	1.201	–0.023	0.054	–0.425
S&LDT2W vs. SSDT2W	0.148	0.098	1.513	0.041	0.063	0.653	0.002	0.058	0.027
S&LDT2W vs. TDT2W	0.014	0.091	0.158	0.039	0.059	0.662	–0.010	0.054	–0.180
SSDT2NW vs. TDT2NW	–0.081	0.053	–1.527	0.012	0.034	0.366	–0.007	0.031	–0.240
SSDT2W vs. TDT2W	–0.134	0.053	–2.541	–0.002	0.034	–0.071	–0.011	0.031	–0.360
CAST4NW vs. SLDT4NW	–0.155	0.123	–1.260	–0.409	0.079	–5.151*	0.102	0.072	1.413
CAST4NW vs. SSDT4NW	–0.171	0.098	–1.749	–0.574	0.063	–9.062*	0.110	0.058	1.904
CAST4NW vs. TDT4NW	–0.256	0.091	–2.811	–0.477	0.059	–8.118*	0.116	0.054	2.165
CAST4W vs. SLDT4W	–0.254	0.123	–2.071	–0.137	0.079	–1.720	0.136	0.072	1.882
CAST4W vs. SSDT4W	–0.152	0.098	–1.554	–0.431	0.063	–6.797*	0.137	0.058	2.367
CAST4W vs. TDT4W	–0.279	0.091	–3.072	–0.349	0.059	–5.939*	0.138	0.054	2.579
S&LDT4NW vs. SSDT4NW	–0.017	0.098	–0.170	–0.165	0.063	–2.604	0.008	0.058	0.133
S&LDT4NW vs. TDT4NW	–0.101	0.091	–1.108	–0.068	0.059	–1.156	0.014	0.054	0.255

(*table continues*)

Table 4. *(Continued).*

Variable	Pitch height			Pitch slope			Pitch curvature		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
S&LDT4W vs. SSDT4W	0.102	0.098	1.041	−0.068	0.059	−1.156	0.000	0.058	0.008
S&LDT4W vs. TDT4W	−0.025	0.091	−0.273	−0.212	0.059	−3.615*	0.002	0.054	0.036
SSDT4NW vs. TDT4NW	−0.084	0.053	−1.596	0.097	0.034	2.849	0.006	0.031	0.193
SSDT4W vs. TDT4W	−0.127	0.053	−2.408	0.082	0.034	2.395	0.001	0.031	0.047

Note. *SE* = standard error; CAS = children with childhood apraxia of speech; S&LD = children with non-CAS speech sound disorder plus language disorder; SSD = children with non-CAS speech sound disorder; TD = children with typical speech-language development; T1 = high-level tone; T2 = high-rising tone; T4 = low-falling tone; W = word; NW = nonword.

*$p < .05$.

High-level tone (T1). For T1 syllables, the SSD group produced pitch heights that were statistically significantly higher on the word compared to the nonword items. No statistically significant differences were found for any of the other relevant within-group comparisons.

High-rising tone (T2). For T2 syllables, the CAS and SSD groups produced pitch heights that were statistically significantly higher on the word compared to the nonword items. No statistically significant differences were found for any of the other relevant within-group comparisons.

Low-falling tone (T4). For T4 syllables, the S&LD and TD groups produced pitch heights that were statistically significantly lower on the word compared to the nonword items. No statistically significant differences were found for any of the other relevant within-group comparisons.

Between-Group Comparisons

Regarding the between-group comparisons on word and nonword syllables with three Cantonese tones, relatively more statistically significant differences were found between the CAS and each of the control groups than in the comparisons among the control groups on T2 pitch curvatures and T4 pitch slopes.

High-level tone (T1). For T1 syllables, the CAS group produced pitch heights that were statistically significantly lower on the word items compared to the SSD group. No statistically significant differences were found for any of the other relevant between-group comparisons.

High-rising tone (T2). For T2 syllables, the CAS group produced pitch slopes that were statistically significantly shallower on the nonword items compared to those of the TD group and pitch curvatures that were statistically significantly less concave on both the word and nonword items compared to those of all of the control groups. No statistically significant differences were found for any of the other relevant between-group comparisons.

Low-falling tone (T4). For T4 syllables, the CAS group produced pitch slopes that were statistically significantly shallower on both the word and nonword items compared to those of all of the control groups, except that there were no significant differences with respect to pitch slopes between the CAS and S&LD groups on the word items. The S&LD group produced pitch slopes that were statistically significantly shallower on word items compared to those of the TD group. No statistically significant differences were found for any of the other relevant between-group comparisons.

Effects of Syllable Position

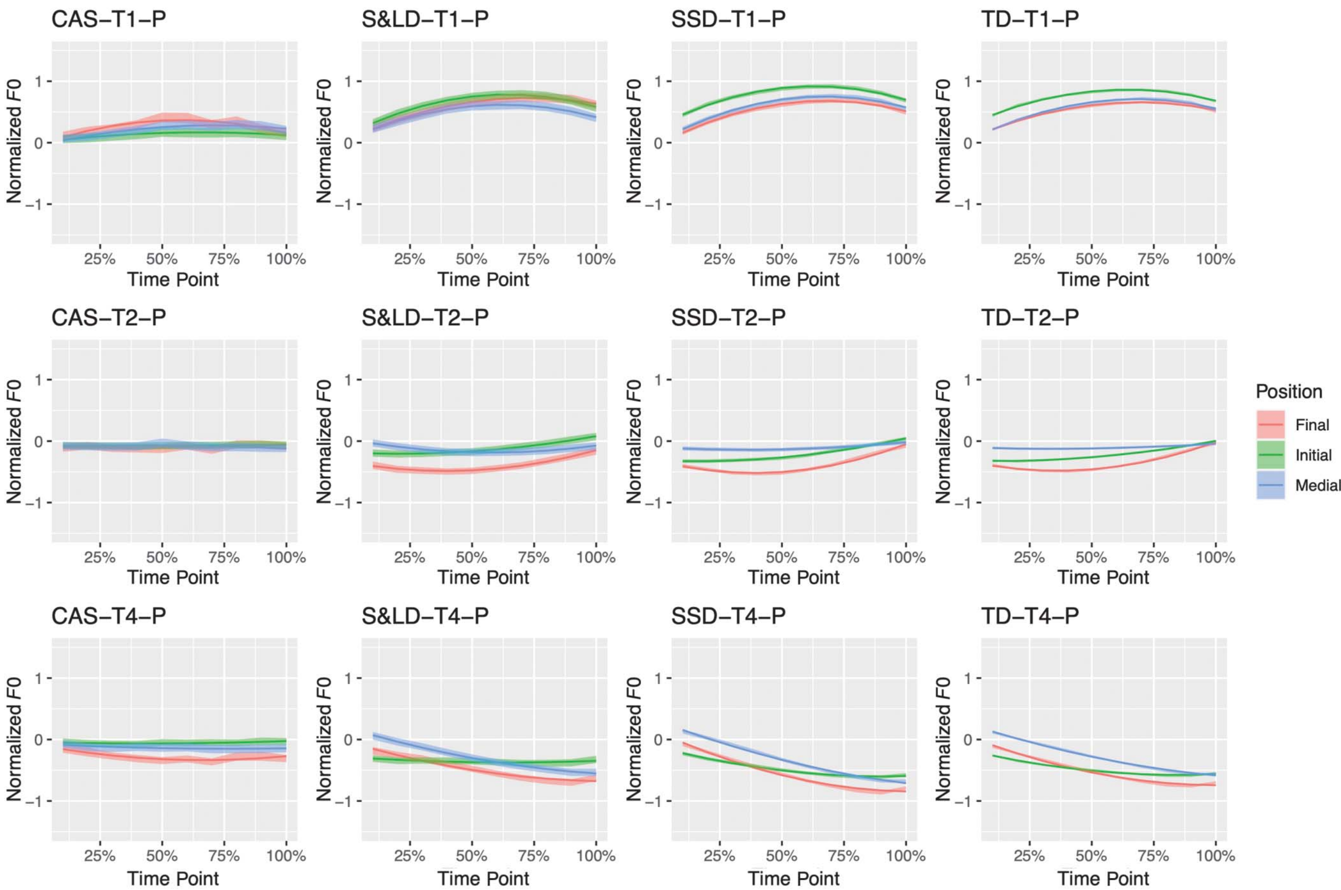
The model fits for normalized *F*0 values by syllable position are shown in Figure 4. Table 5 shows the fixed-effect parameter estimates and their standard errors along with the corresponding *z* values in pairwise comparisons. Only the comparisons of interest are presented. The effect of syllable position on the intercept, $\chi^2(3) = 270{,}872.50$, $p < .05$, and linear, $\chi^2(3) = 27{,}060.40$, $p < .05$, and quadratic, $\chi^2(3) = 7{,}847.70$, $p < .05$, terms did improve model fits, indicating that the *F*0 values of syllables in the initial, medial, and final positions were different regarding pitch heights, slopes, and curvatures.

Within-Group Comparisons

Regarding the within-group comparisons on the three syllable positions on all three of the Cantonese tones, 8, 13, 18, and 23 statistically significant differences were found in the CAS, S&LD, SSD, and TD groups, respectively, on pitch heights, slopes, and curvatures.

High-level tone (T1). For T1 syllables, the CAS group produced the statistically significantly highest pitch heights in the final position, followed by the medial and initial positions, and pitch curvatures that were statistically significantly different between the final and initial positions. The S&LD group produced the statistically significantly highest pitch heights in the initial position, followed by the final and medial positions, and pitch slopes that were statistically significantly different between the final and medial positions. The SSD and TD groups

Figure 4. Growth curve model fits (lines) and observed data (ribbon) for the effect of syllable position on normalized fundamental frequency (*F*0) values on the high-level (T1), high-rising (T2), and low-falling (T4) tones for children with childhood apraxia of speech (CAS), children with non-CAS speech sound disorder plus language disorder (S&LD), children with non-CAS speech sound disorder only (SSD), and children with typical speech-language development (TD).



produced the statistically significantly highest pitch heights in the initial position, followed by the medial and final positions, and pitch slopes that were statistically significantly different between the final versus initial and initial versus medial positions. The TD group produced pitch curvatures that were statistically significant different between the final and medial positions. No statistically significant differences were found for any of the other relevant within-group comparisons.

High-rising tone (T2). For T2 syllables, the CAS group produced comparable pitch heights, slopes, and curvatures in the different positions. The S&LD group had pitch heights that were statistically significantly different between the final versus initial and final versus medial syllables, and pitch slopes that were statistically significantly different between the final versus medial and initial versus medial syllables. The SSD group produced pitch heights that were statistically significantly different in the different positions, pitch slopes that were statistically significantly different between final versus medial and initial versus medial syllables, and pitch curvatures that were statistically significantly different between the final versus initial and final versus medial syllables. The TD group produced pitch heights, slopes, and curvatures that were statistically significantly different between the syllables in the different positions. No statistically significant differences were found for any of the other relevant within-group comparisons.

Low-falling tone (T4). For T4 syllables, the CAS group produced the statistically significantly highest pitch height in the initial position, followed by the medial and final positions and pitch slopes that were statistically significantly different between the initial versus medial syllables. All of the control groups produced the statistically significantly highest pitch heights in the medial position, followed by the initial and final positions. The S&LD group produced pitch slopes that were statistically significantly different between the final versus initial and initial

Table 5. Parameter estimates for analysis of effect of syllable position on normalized fundamental frequency values.

Variable	Pitch height			Pitch slope			Pitch curvature		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
Within-group comparison									
CAST1F vs. CAST1I	−0.136	0.014	−9.938*	0.010	0.043	0.222	0.187	0.043	4.333*
CAST1F vs. CAST1M	−0.056	0.014	−4.077*	0.123	0.043	2.852	0.126	0.043	2.917
CAST1I vs. CAST1M	0.080	0.014	5.860*	0.114	0.043	2.630	−0.061	0.043	−1.416
CAST2F vs. CAST2I	0.025	0.014	1.853	−0.018	0.043	−0.421	−0.020	0.043	−0.460
CAST2F vs. CAST2M	0.010	0.014	0.719	−0.072	0.043	−1.667	−0.064	0.043	−1.469
CAST2I vs. CAST2M	−0.016	0.014	−1.133	−0.054	0.043	−1.246	−0.044	0.043	−1.009
CAST4F vs. CAST4I	0.228	0.014	16.666*	0.143	0.043	3.295	−0.102	0.043	−2.364
CAST4F vs. CAST4M	0.149	0.014	10.892*	0.061	0.043	1.400	−0.099	0.043	−2.293
CAST4I vs. CAST4M	−0.079	0.014	−5.774*	−0.543	0.047	−11.546*	0.049	0.046	1.059
S&LDT1F vs. S&LDT1I	0.055	0.014	4.043*	−0.141	0.043	−3.251	−0.073	0.043	−1.677
S&LDT1F vs. S&LDT1M	−0.096	0.014	−6.986*	−0.207	0.043	−4.783*	−0.044	0.043	−1.024
S&LDT1I vs. S&LDT1M	−0.151	0.014	−11.028*	−0.066	0.043	−1.532	0.028	0.043	0.652
S&LDT2F vs. S&LDT2I	0.274	0.014	19.990*	0.024	0.043	0.560	−0.109	0.043	−2.518
S&LDT2F vs. S&LDT2M	0.256	0.014	18.709*	−0.298	0.043	−6.882*	−0.064	0.043	−1.484
S&LDT2I vs. S&LDT2M	−0.018	0.014	−1.281	−0.322	0.043	−7.442*	−0.054	0.043	−1.255
S&LDT4F vs. S&LDT4I	0.127	0.014	9.263*	0.490	0.043	11.320*	−0.077	0.043	−1.776
S&LDT4F vs. S&LDT4M	0.183	0.014	13.370*	−0.100	0.043	−2.305	−0.017	0.043	−0.399
S&LDT4I vs. S&LDT4M	0.056	0.014	4.107*	−0.590	0.043	−13.625*	0.060	0.043	1.377
SSDT1F vs. SSDT1I	0.246	0.007	34.378*	−0.111	0.023	−4.894*	−0.013	0.023	−0.573
SSDT1F vs. SSDT1M	0.066	0.007	9.282*	0.005	0.023	0.199	−0.013	0.023	−0.561
SSDT1I vs. SSDT1M	−0.179	0.007	−25.097*	0.115	0.023	5.094*	0.000	0.023	0.013
SSDT2F vs. SSDT2I	0.178	0.007	24.966*	0.010	0.023	0.444	−0.169	0.023	−7.485*
SSDT2F vs. SSDT2M	0.277	0.007	38.692*	−0.262	0.023	−11.592*	−0.221	0.023	−9.799*
SSDT2I vs. SSDT2M	0.098	0.007	13.726*	−0.272	0.023	−12.037*	−0.052	0.023	−2.314
SSDT4F vs. SSDT4I	0.076	0.007	10.631*	0.427	0.023	18.889*	−0.071	0.023	−3.164
SSDT4F vs. SSDT4M	0.215	0.007	30.088*	−0.072	0.023	−3.174	−0.087	0.023	−3.857*
SSDT4I vs. SSDT4M	0.139	0.007	19.457*	−0.499	0.023	−22.063*	−0.016	0.023	−0.693
TDT1F vs. TDT1I	0.209	0.004	49.496*	−0.085	0.013	−6.364*	−0.027	0.013	−2.046
TDT1F vs. TDT1M	0.035	0.004	8.296*	0.030	0.013	2.256	−0.044	0.013	−3.301*
TDT1I vs. TDT1M	−0.174	0.004	−41.200*	0.115	0.013	8.620*	−0.017	0.013	−1.255
TDT2F vs. TDT2I	0.135	0.004	32.005*	−0.062	0.013	−4.613*	−0.166	0.013	−12.425*
TDT2F vs. TDT2M	0.245	0.004	58.098*	−0.319	0.013	−23.856*	−0.219	0.013	−16.401*
TDT2I vs. TDT2M	0.110	0.004	26.093*	−0.257	0.013	−19.242*	−0.053	0.013	−3.977*
TDT4F vs. TDT4I	0.031	0.004	7.437*	0.353	0.013	26.469*	−0.049	0.013	−3.685*
TDT4F vs. TDT4M	0.228	0.004	54.007*	−0.065	0.013	−4.863*	−0.073	0.013	−5.429*
TDT4I vs. TDT4M	0.197	0.004	46.570*	−0.418	0.013	−31.332*	−0.023	0.013	−1.744
Between-group comparison									
CAST1F vs. S&LDT1F	0.317	0.020	15.556*	−0.181	0.043	−4.189*	0.187	0.043	4.333*
CAST1I vs. S&LDT1I	0.508	0.020	24.942*	0.168	0.051	3.314	0.126	0.043	2.917
CAST1M vs. S&LDT1M	0.277	0.020	13.603*	0.012	0.064	0.191	−0.061	0.043	−1.416
CAST1F vs. SSDT1F	0.259	0.016	15.940*	0.342	0.064	5.391*	−0.020	0.043	−0.460
CAST1I vs. SSDT1I	0.641	0.016	39.413*	0.192	0.064	3.025	−0.064	0.043	−1.469
CAST1M vs. SSDT1M	0.381	0.016	23.451*	0.169	0.051	3.340	−0.044	0.043	−1.009
CAST1F vs. TDT1F	0.260	0.015	17.215*	0.253	0.047	5.382*	−0.102	0.043	−2.364
CAST1I vs. TDT1I	0.605	0.015	40.089*	0.158	0.047	3.370	−0.099	0.043	−2.293
CAST1M vs. TDT1M	0.350	0.015	23.237*	0.160	0.047	3.398	0.049	0.046	1.059
S&LDT1F vs. SSDT1F	−0.058	0.016	−3.561*	−0.054	0.051	−1.072	−0.027	0.013	−2.046
S&LDT1F vs. TDT1F	−0.057	0.015	−3.807*	−0.089	0.047	−1.903	−0.044	0.013	−3.301*

(*table continues*)

Table 5. *(Continued).*

Variable	Pitch height			Pitch slope			Pitch curvature		
	Estimate	*SE*	*z*	Estimate	*SE*	*z*	Estimate	*SE*	*z*
S&LDT1I vs. SSDT1I	0.132	0.016	8.147*	–0.024	0.051	–0.478	–0.017	0.013	–1.255
S&LDT1I vs. TDT1I	0.096	0.015	6.384*	–0.034	0.047	–0.718	–0.166	0.013	–12.425*
S&LDT1M vs. SSDT1M	0.104	0.016	6.398*	0.157	0.051	3.101	–0.219	0.013	–16.401*
S&LDT1M vs. TDT1M	0.073	0.015	4.854*	0.148	0.047	3.140	–0.053	0.013	–3.977*
SSDT1F vs. TDT1F	0.000	0.009	0.055	–0.035	0.027	–1.290	0.183	0.063	2.924
SSDT1I vs. TDT1I	–0.036	0.009	–4.137*	–0.010	0.027	–0.349	0.261	0.050	5.215*
SSDT1M vs. TDT1M	–0.031	0.009	–3.524	–0.010	0.027	–0.349	0.234	0.046	5.040*
CAST2F vs. S&LDT2F	–0.292	0.020	–14.320*	0.222	0.064	3.500	–0.073	0.043	–1.677
CAST2F vs. SSDT2F	–0.286	0.016	–17.580*	0.329	0.051	6.482*	–0.044	0.043	–1.024
CAST2F vs. TDT2F	–0.250	0.015	–16.561*	0.355	0.047	7.543*	0.028	0.043	0.652
CAST2I vs. S&LDT2I	–0.044	0.020	–2.143	0.265	0.064	4.168*	–0.109	0.043	–2.518
CAST2I vs. SSDT2I	–0.133	0.016	–8.167*	0.357	0.051	7.039*	–0.064	0.043	–1.484
CAST2I vs. TDT2I	–0.140	0.015	–9.281*	0.311	0.047	6.620*	–0.054	0.043	–1.255
CAST2M vs. S&LDT2M	–0.046	0.020	–2.242	–0.003	0.064	–0.052	–0.077	0.043	–1.776
CAST2M vs. SSDT2M	–0.019	0.016	–1.180	0.139	0.051	2.736	–0.017	0.043	–0.399
CAST2M vs. TDT2M	–0.014	0.015	–0.948	0.108	0.047	2.302	0.060	0.043	1.377
S&LDT2F vs. SSDT2F	0.006	0.016	0.370	0.106	0.051	2.094	–0.049	0.013	–3.685*
S&LDT2F vs. TDT2F	0.042	0.015	2.789	0.132	0.047	2.814	–0.073	0.013	–5.429*
S&LDT2I vs. SSDT2I	–0.089	0.016	–5.480*	0.092	0.051	1.814	–0.023	0.013	–1.744
S&LDT2I vs. TDT2I	–0.096	0.015	–6.386*	0.046	0.047	0.987	0.307	0.043	7.094*
S&LDT2M vs. SSDT2M	0.027	0.016	1.630	0.142	0.051	2.801	–0.232	0.046	–4.993*
S&LDT2M vs. TDT2M	0.031	0.015	2.082	0.111	0.047	2.371	–0.183	0.063	–2.911
SSDT2F vs. TDT2F	0.036	0.009	4.123*	0.026	0.027	0.959	0.094	0.063	1.505
SSDT2I vs. TDT2I	–0.007	0.009	–0.825	–0.046	0.027	–1.670	0.112	0.050	2.234
SSDT2M vs. TDT2M	0.005	0.009	0.560	–0.030	0.027	–1.118	0.088	0.046	1.896
CAST4F vs. S&LDT4F	–0.199	0.020	–9.767*	–0.414	0.064	–6.521*	–0.013	0.023	–0.573
CAST4F vs. SSDT4F	–0.273	0.016	–16.764*	–0.689	0.051	–13.587*	–0.013	0.023	–0.561
CAST4F vs. TDT4F	–0.230	0.015	–15.226*	–0.082	0.043	–1.895	0.000	0.023	0.013
CAST4I vs. S&LDT4I	–0.300	0.020	–14.737*	–0.067	0.064	–1.055	–0.169	0.023	–7.485*
CAST4I vs. SSDT4I	–0.425	0.016	–26.118*	–0.404	0.051	–7.978*	–0.221	0.023	–9.799*
CAST4I vs. TDT4I	–0.426	0.015	–28.264*	–0.332	0.047	–7.060*	–0.052	0.023	–2.314
CAST4M vs. S&LDT4M	–0.165	0.020	–8.103*	–0.575	0.064	–9.045*	–0.071	0.023	–3.164
CAST4M vs. SSDT4M	–0.207	0.016	–12.707*	–0.821	0.051	–16.198*	–0.087	0.023	–3.857*
CAST4M vs. TDT4M	–0.151	0.015	–9.987*	–0.668	0.047	–14.216*	–0.016	0.023	–0.693
S&LDT4F vs. SSDT4F	–0.074	0.016	–4.521*	–0.274	0.051	–5.412*	–0.085	0.050	–1.696
S&LDT4F vs. TDT4F	–0.031	0.015	–2.029	–0.129	0.047	–2.734	–0.272	0.063	–4.337*
S&LDT4I vs. SSDT4I	–0.124	0.016	–7.645*	–0.337	0.051	–6.655*	–0.224	0.050	–4.470*
S&LDT4I vs. TDT4I	–0.126	0.015	–8.350*	–0.265	0.047	–5.635*	–0.017	0.046	–0.368
S&LDT4M vs. SSDT4M	–0.041	0.016	–2.549	–0.246	0.051	–4.859*	–0.285	0.050	–5.700*
S&LDT4M vs. TDT4M	0.015	0.015	0.963	–0.094	0.047	–1.994	–0.187	0.046	–4.034*
SSDT4F vs. TDT4F	0.043	0.009	4.909*	0.146	0.027	5.349*	0.183	0.063	2.913
SSDT4I vs. TDT4I	–0.002	0.009	–0.188	0.072	0.027	2.656	0.103	0.050	2.061
SSDT4M vs. TDT4M	0.056	0.009	6.402*	0.153	0.027	5.598*	0.079	0.046	1.693

Note. *SE* = standard error; CAS = children with childhood apraxia of speech; S&LD = children with non-CAS speech sound disorder plus language disorder; SSD = children with non-CAS speech sound disorder; TD = children with typical speech-language development; T1 = high-level tone; T2 = high-rising tone; T4 = low-falling tone; I = initial position; M = medial position; F = final position.

*$p < .05$.

versus medial syllables. The SSD groups produced pitch slopes that were statistically significantly different between the final versus initial and initial versus medial syllables and pitch curvatures that were statistically significantly different between final versus medial syllables. The TD group showed pitch slopes that were statistically significantly different in the different positions and pitch curvatures that were statistically significantly different between the final versus initial and final versus medial positions. No statistically significant differences were found for any of the other relevant within-group comparisons.

Between-Group Comparisons

Regarding the between-group comparisons on the three syllable positions with three Cantonese tones, more statistically significant differences were found between the CAS and each of the control groups, compared to the comparisons among the control groups, with respect to pitch heights and slopes. More statistically significant differences were found between the S&LD group and the other three groups (including the CAS group) than among the other groups with respect to pitch curvatures.

High-level tone (T1). For T1 syllables, the CAS group produced pitch heights that were statistically significantly different in different positions compared to all of the control groups, pitch slopes that were statistically significantly different in the final position compared to all of the control groups, and pitch slopes that were statistically significantly different in the final position compared to the S&LD group. The S&LD group produced pitch heights that were statistically significantly different in different positions compared to the SSD and TD groups and pitch curvatures that were statistically significantly different in the final and medial positions compared to the TD group and in the medial and initial positions compared to the SSD group. The SSD group produced pitch heights that were statistically significantly different in the initial position compared to the TD group and pitch curvatures that were statistically significantly different in the initial and medial positions compared to the TD group. No statistically significant differences were found for any of the other relevant between-group comparisons.

High-rising tone (T2). For T2 syllables, the CAS group produced pitch heights that were statistically significantly different in the final position compared to all of the three control groups and initial position compared to the SSD and TD groups and pitch slopes that were statistically different in the final position compared to the SSD and TD groups and initial position compared to all the three control groups. The S&LD group produced pitch heights that were statistically significantly different in the initial position compared to the SSD and TD groups and pitch curvatures that were statistically significantly different in the final and medial positions compared to the SSD group and final and initial positions compared to the TD group. The SSD group produced pitch heights that were statistically significantly different in the final position compared to the TD group. No statistically significant differences were found for any of the other relevant between-group comparisons.

Low-falling tone (T4). For T4 syllables, the CAS group produced pitch heights that were statistically significantly different in different positions compared to all of the three control groups; pitch slopes that were statistically significantly different in final position compared to the S&LD and SSD groups, in the initial position compared to the SSD and TD group, and in the medial position compared to all of the control groups and pitch curvatures that were statistically significantly different in the initial position compared to the S&LD and SSD groups and medial position compared to the SSD group. The S&LD group produced pitch heights that were statistically significantly different in the final and initial positions compared to the SSD group and in the initial position compared to the TD group, pitch slopes that were statistically significantly different in all three of the positions compared to the SSD group and in the initial position compared to the TD group, and pitch curvatures that were statistically significantly different in the initial and medial positions compared to the SSD group and in the final and medial positions compared to the TD group. The SSD group produced pitch heights and slopes that were statistically significantly different in the final and medial positions compared to the TD group. No statistically significant differences were found for any of the other relevant between-group comparisons.

Discussion

In this study, we aimed to examine *F*0 variations and the effects of syllable structure, lexical status, and syllable position on *F*0 in Cantonese-speaking preschool children with CAS, compared with those with non-CAS speech and language disorders, those with non-CAS speech sound disorders only, and those with typical speech-language development. Four hypotheses were tested.

The first hypothesis, that the CAS group would show significant differences in *F*0 production on the three Cantonese tones compared with the three control groups while comparable performance would be observed among the control groups, was partially supported. The CAS group showed significantly lower pitch heights on T1 (high-level) syllables compared to the SSD and TD groups, but not to the S&LD group; significantly shallower pitch slopes on T2 (high-rising) and T4 (low-falling) syllables compared to the SSD and TD groups, but not to the S&LD group; and significant reduction in concave curvature on T2 (high rising) syllables compared to all of the control groups. The three control groups showed comparable *F*0 productions among each other. The results corroborate and extend the initial discoveries outlined by Wong et al. (2021) by confirming that Cantonese-speaking children with CAS have difficulty with *F*0/pitch variation skills, resulting in reduced tone contrasts. These children exhibit a tendency to generate lower pitch heights when a high-level tone is anticipated and minimal pitch fluctuation within a syllable when a contour tone is expected. These difficulties may be attributed to impaired planning and programming skills that impede motor control of laryngeal musculature (Liang & Du, 2018). In particular, children with CAS may have difficulty in planning and programming the laryngeal movements that involve contraction of the cricothyroid to produce a rising pitch and reduction of subglottal pressure or intrinsic laryngeal tension to produce a falling pitch (Snow, 2007).

These results are inconsistent with those of previous studies that examined the *F*0 of lexical stress errors in English speakers with CAS and found no acoustic differences between children with CAS and those with non-CAS speech sound disorders or typical speech-language development (Kopera & Grigos, 2020; Munson et al., 2003). This inconsistency may be due to the differing linguistic features of Cantonese versus English. Specifically, pitch variation plays different roles in the two languages. Cantonese, a tonal language, requires pitch variations within syllables to express different lexical tones, which in turn convey different lexical meanings, while maintaining constant syllable durations. In contrast, English requires simultaneous variations of pitch, loudness, and syllable duration to express lexical stress patterns. Another reason for the variant results may be that the methodologies of the previous studies may have failed to detect the actual pitch variation difficulties in children with CAS, which were revealed by using GCA in this study. This study contributes to the literature that *F*0 values, as they change within the syllable, are effective in differentiating Cantonese speakers with and without CAS.

The second hypothesis, that the CAS group would show significant within-group differences in *F*0 production between V versus CV syllables with three Cantonese tones and significantly different pitches compared with the three control groups, was partially supported. The results of within-group comparisons deviated somewhat from the initial prediction, while the findings from the between-group comparisons were consistent with the predictions. Regarding within-group comparisons, the CAS group displayed fewer significant differences in pitch slopes and curvatures between V and CV syllables for all three of the Cantonese tones when considering the number of variations between the CAS and control groups. The fewer *F*0 contrasts between V and CV syllables could be explained by their difficulty in simultaneous management of syllable structure and tone. English (nontonal language) speakers with CAS have increased difficulty with increased linguistic complexity (ASHA, 2007). This implies that more complex syllables pose a greater challenge for these children. In tonal language speakers, the presence of tones associated with vowels presents additional motoric challenges. In this study of Cantonese speakers, the CAS group produced similar *F*0 patterns between syllables that had different syllable structures. It is likely that when a contour tone, such as a rising or falling tone, is required, the motoric demands escalate, and the contrast is even harder to achieve, resulting in even more similar productions between syllables with different syllable structures.

Regarding the between-group comparisons on V and CV syllables with three Cantonese tones, more statistically significances were found between the CAS and each of the control groups, compared to the comparisons among the control groups, in pitch heights, slopes, and curvatures. The results were inconsistent with the perceptual findings reported by Wong et al. (2021), who found no group differences in perceived tone accuracy on syllables with V versus CV structures. The current study demonstrated that there were actual acoustic differences on *F*0 in different syllable structures between the children with and without CAS. Both the within-group and between-group comparisons suggested that children with CAS may be unable to manage *F*0 and syllable structures simultaneously, resulting in more similar productions. It appears that Cantonese-speaking children with CAS are unable to produce the tone contrasts that can be found in children without CAS, suggesting a potential measure for differential diagnosis. Apart from the established knowledge that children with CAS have difficulty in producing correct phonemes in complex syllable structures (ASHA, 2007; Canault et al., 2021), the current study supports the notion that these children also encounter difficulties in planning other aspects of syllables, especially their prosodic components (Nijland et al., 2003). The study also highlights the challenge of pitch variation in children with CAS, thereby broadening the understanding of their speech production difficulties.

The third hypothesis, that all of the groups would show significant within-group differences in *F*0 production between the word and nonword items with the three Cantonese tones while the CAS group would show significantly different *F*0 production on both the word and nonword items compared to the three control groups, was not supported. In the within-group comparisons, each of the groups showed significant differences in pitch heights only when lexical status was considered and the significant differences were fewer than those found in the analyses regarding syllable structure. In the between-group comparisons, again, the total number of significant differences was less than that found in the syllable structure analyses. Although relatively more statistically significant differences were found between the CAS and each of the control groups than among the control groups on T2 pitch curvatures and T4 pitch slopes, these results can be explained by the generally reduced levels of *F*0 contrast found in the CAS group, especially on contour tones. With respect to the T1 word and nonword syllables, only one statistically significant difference was obtained in pitch height (i.e., participants with CAS producing T1 word syllable [CAST1W] vs. participants with non-CAS speech sound disorder only producing T1 word syllable [SSDT1W]). We conclude that the groups did not have significantly different *F*0 values on the word versus the nonword items. These outcomes align with the findings presented by Wong et al. (2021), who perceptually identified no discrepancies in terms of tone accuracy between word and nonword items produced by Cantonese-speaking children with CAS.

Although lexical meanings prime the production of word items in Cantonese speakers (Zhao & Jurafsky, 2009), these priming effects work differently in children with and without CAS. It was observed that the CAS and SSD groups produced T2 (high-rising) word syllables with significantly higher pitch heights than T2 nonword syllables, and the SSD group produced T1 (high-level) word syllables with significantly higher pitch heights than T1 nonword syllables. The S&LD and TD groups produced T4 (low-falling) word syllables with significantly lower pitch heights than T4 nonword syllables. The groups apparently increased their vocal fold tension to produce higher pitch heights for high tones and reduced the vocal fold tension to produce lower pitch heights for low tones in words, thus amplifying the distinctions between the tones in words.

However, it should be noted that the priming effects did not override the difficulty of pitch variation skills in Cantonese-speaking children with CAS. This result is different from the responses of Cantonese-speaking adults, who hyperarticulated low-frequency items, resulting in higher pitch heights on those items (Zhao & Jurafsky, 2009). This could be explained by methodological differences. The children in this study may have hyperarticulated word items compared with nonword items due to the requirement of repeating stimuli five times consecutively during the TST. That is, children may exert more effort when processing word items in order to achieve clear boundaries between the repetitions of the words. In contrast, the boundaries between nonword items may have been less distinct due to the absence of clear mental representations for these items, leading to fewer cognitive demands and potentially less vocal effort or tension required to produce repetitions. Nevertheless, additional research is recommended to identify the factors that contribute to the facilitation observed on certain tones.

The last hypothesis, that all of the groups would show significant within-group differences in *F*0 among the syllables in different positions with the three Cantonese tones while the CAS group would show a greater number of significant differences in *F*0 on the syllables in all of the three positions compared to the three control groups, was partially supported. The results of the within-group comparisons supported the initial prediction, while the findings from the between-group comparisons did not. The results of the within-group comparisons showed that there were actual acoustic differences among the syllables in different positions, regardless of the clinical diagnosis. This is consistent with the findings reported by Li and Chen (2016) and Fon and Hsu (2004) and extends the findings of the latter pair of researchers from Mandarin-speaking adults to Cantonese-speaking children. However, the CAS group produced fewer significant differences compared to each of the control groups, which was unexpected, given that children with CAS have more variable articulatory movements (Grigos et al., 2015). This may be explained by their reduced *F*0/tone contrasts overall and the fact that tone production involves more complex articulatory movements, compared to nontonal phoneme productions. Since tones are associated with vowels, tone productions require not only complex oral articulatory movements but also precise control of the laryngeal muscles. Children with CAS may encounter challenges in planning and programming these complex oral and laryngeal articulatory movements, resulting in more similar tone productions.

In addition, when considering variations among syllables in different positions, the CAS group exhibited a distinct pattern in contrast to those of the three control groups. Since the CAS group showed difficulty in producing the rising and falling contours (i.e., had reduced tone contrasts), the following discussion is based on the relatively easiest tone among the three (i.e., high level, T1). The CAS group produced the highest pitch levels for T1 (high-level) syllables in the final position, followed by the medial and initial positions. The three control groups produced the highest T1 pitch levels in the initial position, followed by the medial and final positions. Given that a higher pitch height may reflect greater effort at pitch production (Zhao & Jurafsky, 2009), we speculate that there are distinct approaches to pitch production in the sequencing context among Cantonese-speaking children with CAS versus those from other diagnostic groups. Based on the computational model of speech production (Guenther, 2016), children with CAS exhibit deficits in the feedforward control system, leading to a greater reliance on the feedback control system, especially the auditory feedback control system (Iuzzini-Seigel et al., 2015). It is possible that Cantonese speakers with CAS have deviant speech sound maps containing variable and incomplete tonal information alongside their segmental information. This can be observed based on the result that the participants in the CAS group produced more variable within-group *F*0 values, compared to the control groups. In contrast, children without CAS may have complete tonal information in their speech sound maps. With respect to the production of tones and tone sequences, it is postulated that children without CAS generate the relevant motor commands for tones from their speech sound maps. In contrast, children with CAS struggle to generate precise motor commands for tones from their speech sound maps, leading to inaccurate productions. As the sequence progresses, children without CAS may reduce their efforts in producing the syllables in the medial and final positions, resulting in lowering the pitch heights. However, because children with CAS tend to rely on auditory feedback from their own productions and compare them to their own inadequate speech sound map tonal information, they may recognize, belatedly, that their productions are deviant. Consequently, children with CAS may keep making effort on the productions, resulting in higher pitch heights in the later positions. However, it is important to emphasize that these explanations are speculative, and further research is necessary to validate these assumptions with empirical evidence.

Clinical Implications

Overall, this study demonstrated that tone production in sequencing contexts, measured using the TST in this study, could differentiate Cantonese-speaking children with CAS from those without CAS. The results imply that the impairment profiles in CAS are likely language specific. When assessing individuals with CAS, it is crucial to consider these language-specific features of the impairment. Language-specific assessments and interventions are vital to address the unique challenges and needs associated with CAS in specific languages. Given that there are no objective tools reported for the diagnosis of CAS in Cantonese speakers (Wong, Wong, & Velleman, 2023b), the TST adopted in this study could be further developed as a clinical tool by examining the effectiveness of the individual items in differentiating children with and without CAS.

Limitations and Future Investigations

This study was limited in several ways. First, although this study included six Cantonese-speaking children with CAS, which was more than those that were included in previous research focused on Cantonese speakers with the same disorder (Wong et al., 2021; Wong, Wong, & Velleman, 2023a; Wong, Wong, Velleman, Tong, & Lee, 2023), the findings are limited by this small sample size. Second, the cognitive abilities of the participants were not tested in this study. Any undetected cognitive impairments might have influenced the results. In addition, the relatively large number of missing values in the CAS group, compared with the control groups, may have influenced the findings, despite the fact that a robust imputation method was applied. Moreover, regarding acoustic analyses, the procedure adopted in this study did not address the potential presence of pitch-tracking errors or glottal fry, which may have influenced the *F*0 values. Finally, in the interest of maintaining a reasonable number of data points for GCA, the inclusion of offset (100%) *F*0 values, which could have been influenced by the subsequent syllables during syllable repetition, might have introduced limitations in generalizing the results.

In addition to the development of a clinical tool based on the TST, future investigations should focus on applying GCA to examine *F*0 changes across time in English speakers with CAS. Given that lexical stress involves simultaneous variations of pitch, loudness, and duration, GCA could provide a statistical method to capture pitch and/or intensity values across syllables, resulting in a deeper understanding of pitch and loudness variation skills. Furthermore, the TST provides a medium to investigate pitch variation by keeping syllable duration constant. The application of the TST to English or to speakers of other stress-timed languages may provide a method for understanding pitch variation skills in isolation and whether they play a role in those speakers. Finally, the TST is based on repetitions of tones and tone sequences. Future investigations can be focused on evaluating the consistency of tone in syllable repetitions and tone sequences in sequence repetitions.

Conclusions

This study used GCA to show that Cantonese-speaking children with CAS have difficulty with pitch variation. Their deficits were further investigated by examining the effects on tone production of three linguistic aspects, namely, syllable structure, lexical status, and syllable position. Regarding pitch performance, children with CAS are unable to produce the pitch contrasts that can be found in children without CAS. Thus, testing pitch variation skills may be useful for differentiating children with and without CAS who speak tone languages. Future investigations are recommended in order to develop a quantitative measure for Cantonese speakers with CAS using $F0$ changes within syllables. Applications of GCA and the TST to speakers of other languages with CAS are also suggested to further examine CAS in cross-linguistic investigations.

Data Availability Statement

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This study was supported by the General Research Fund (GRF) of the Hong Kong SAR Government (Project No. 15605821), granted to the first, second, and third authors. The authors would like to express their gratitude to the Hong Kong SAR Government for their support through GRF. Additionally, the authors convey their appreciation to the children and their families for participating in the study. They are also grateful to the postgraduate and undergraduate student helpers as well as research assistants for their assistance in collecting data and annotating the acoustic data, including Ellen Chan, Tina Chan, Oscar Cheng, Ashley Cheung, Twinkie Cheung, Jocelyn Ching, Helen Chung, Edwin Lai, Billy Lee, Jason Lee, Nancy Li, Suki Lo, Jess Loh, Jenna Mak, Carly Ng, Sheldon Sit, Ventus Siu, Dahlia Tang, Florence Wong, Carol Wu, Audrey Yim, Miki Yip, Sicilia Wong, Cherry Yu, and Niki Yu. The authors also would like to extend their appreciation to the anonymous reviewers for their valuable feedback, which greatly contributed to the refinement of this article.

References

Akmam, E. F., Siswantining, T., Soemartojo, S. M., & Sarwinda, D. (2019). *Multiple imputation with predictive mean matching method for numerical missing data*. 2019 3rd International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia.

American Academy of Audiology. (1997). Identification of hearing loss and middle ear dysfunction in children [Position statements]. *Audiology Today, 9*(3), 18–20.

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

Andrus, T., Dubinski, E., Fiscus, J., Gillies, B., Harper, M., Hazen, T. J., Hefright, B., Jarrett, A., Lin, W., Ray, J., Rytting, A., Shen, W., Tzoukermann, E., & Wong, J. (2016). *IARPA Babel Cantonese Language Pack IARPA-babel101b-v0.4c (LDC2016S02)*. Linguistic Data Consortium. https://doi.org/10.35111/4m6a-1x29

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (Version 6.2.14) [Computer program]. http://www.praat.org/

Canault, M., Thai-van, H., & Le Normand, M. T. (2021). Simplification of syllable structure in childhood apraxia of speech: A 2-year follow-up French case study. *Clinical Linguistics & Phonetics, 35*(10), 945–963. https://doi.org/10.1080/02699206.2020.1839971

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.

Child Assessment Service. (2020). *The Hong Kong Test of Preschool Oral Language (Cantonese)*.

Chung, H. C., Wong, E. C. H., Wong, M. N., Velleman, S. L., & Loh, J. (2023). *Performance on syllable repetition tasks by Cantonese-speaking preschool children with childhood apraxia of speech and typical development*. 2023 Symposium of the International Clinical Phonetics and Linguistics Association, Salzburg, Austria.

Fon, J., & Hsu, H.-J. (2004). *Positional and phonotactic effects on the realization of Taiwan Mandarin tone 2*. The 8th International Conference on Spoken Language Processing, Jeju, South Korea. https://www.isca-speech.org/archive/pdfs/interspeech_2004/hsu04_interspeech.pdf [PDF]

Grigos, M. I., Moss, A., & Lu, Y. (2015). Oral articulatory control in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 58*(4), 1103–1118. https://doi.org/10.1044/2015_JSLHR-S-13-0221

Guenther, F. H. (2016). *Neural control of speech*. MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001

Iuzzini-Seigel, J., Allison, K. M., & Stoeckel, R. (2022). A tool for differential diagnosis of childhood apraxia of speech and dysarthria in children: A tutorial. *Language, Speech, and Hearing Services in Schools, 53*(4), 926–946. https://doi.org/10.1044/2022_LSHSS-21-00164

Iuzzini-Seigel, J., Hogan, T. P., & Green, J. R. (2017). Speech inconsistency in children with childhood apraxia of speech, language impairment, and speech delay: Depends on the stimuli. *Journal of Speech, Language, and Hearing Research, 60*(5), 1194–1210. https://doi.org/10.1044/2016_jslhr-s-15-0184

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. https://doi.org/10.1016/j.jcomdis.2015.01.002

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kopera, H. C., & Grigos, M. I. (2020). Lexical stress in childhood apraxia of speech: Acoustic and kinematic findings. *International Journal of Speech-Language Pathology, 22*(1), 12–23. https://doi.org/10.1080/17549507.2019.1568571

Lai, R., & Winterstein, G. (2020). *Cifu: A frequency lexicon of Hong Kong Cantonese*. The Twelfth Language Resources and Evaluation Conference, Marseille, France.

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

Leung, M. T., & Law, S. P. (2001). HKCAC: The Hong Kong Cantonese adult language corpus. *International Journal of Corpus Linguistics, 6*(2), 305–325. https://doi.org/10.1075/ijcl.6.2.06leu

Li, Q., & Chen, Y. (2016). An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics, 54,* 123–150. https://doi.org/10.1016/j.wocn.2015.10.002

Liang, B., & Du, Y. (2018). The functional neuroanatomy of lexical tone perception: An activation likelihood estimation meta-analysis. *Frontiers in Neuroscience, 12,* Article 495. https://doi.org/10.3389/fnins.2018.00495

Luke, K. K., & Wong, M. L. Y. (2015). The Hong Kong Cantonese corpus: Design and uses. *Journal of Chinese Linguistics, 25,* 309–330.

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. https://doi.org/10.1121/1.2363927

Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.

Munson, B., Bjorum, E. M., & Windsor, J. (2003). Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *Journal of Speech, Language, and Hearing Research, 46*(1), 189–202. https://doi.org/10.1044/1092-4388(2003/015)

Murray, E., Iuzzini-Seigel, J., Maas, E., Terband, H., & Ballard, K. J. (2021). Differential diagnosis of childhood apraxia of speech compared to other speech sound disorders: A systematic review. *American Journal of Speech-Language Pathology, 30*(1), 279–300. https://doi.org/10.1044/2020_AJSLP-20-00063

Murray, E., McCabe, P., Heard, R., & Ballard, K. (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/10.1044/2014_JSLHR-S-12-0358

Nijland, L., Maassen, B., Van Der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 17*(1), 1–24. https://doi.org/10.1080/0269920021000050662

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*(3), 271–277. https://doi.org/10.1044/jshd.5203.271

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication, 6*(4), 343–352. https://doi.org/10.1016/0167-6393(87)90009-4

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

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The syllable repetition task (SRT). *Journal of Speech, Language, and Hearing Research, 52*(5), 1189–1212. https://doi.org/10.1044/1092-4388(2009/08-0047)

Shriberg, L. D., Potter, N. L., & Strand, E. A. (2011). Prevalence and phenotype of childhood apraxia of speech in youth with galactosemia. *Journal of Speech, Language, and Hearing Research, 54*(2), 487–519. https://doi.org/10.1044/1092-4388(2010/10-0068)

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeny, J. L., Tilkens, C. M., & Wilson, D. L. (2017). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: I. Development and description of the pause marker. *Journal of Speech, Language, and Hearing Research, 60*(4), S1096–s1117. https://doi.org/10.1044/2016_jslhr-s-15-0296

Snow, D. (2007). Polysyllabic units in the vocalizations of children from 0;6 to 1;11: Intonation-groups, tones and rhythms. *Journal of Child Language, 34*(4), 765–797. https://doi.org/10.1017/S030500090700815X

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., & McCauley, R. J. (2019). *Dynamic Evaluation of Motor Speech Skill*. Brookes.

Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics, 13*(1), 1–23. https://doi.org/10.1080/026992099299211

Wong, E. C. H., Lee, K. Y. S., & Tong, M. C. F. (2019). *Tone sequencing in childhood apraxia of speech: A pilot study*. The 31st World Congress of the International Association of Logopedics and Phoniatrics, Taipei, Taiwan.

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/10.1044/2021_AJSLP-20-00150

Wong, E. C. H., Wong, M. N., Chen, S., & Lin, J. Y. W. (2024). Pitch variation skills in Cantonese speakers with apraxia of speech after stroke: Preliminary findings of acoustic analyses. *Journal of Speech, Language, and Hearing Research, 67*(1), 1–33. https://doi.org/10.1044/2023_JSLHR-23-00242

Wong, E. C. H., Wong, M. N., & Velleman, S. L. (2023a). Assessment and diagnostic standards of apraxia of speech in Chinese-speaking adults and children: A scoping review. *American Journal of Speech-Language Pathology, 32*(1), 316–340. https://doi.org/10.1044/2022_ajslp-21-00355

Wong, E. C. H., Wong, M. N., & Velleman, S. L. (2023b). Clinical practice of childhood apraxia of speech in Hong Kong: A web-based survey study. *PLOS ONE, 18*(4), Article e0284109. https://doi.org/10.1371/journal.pone.0284109

Wong, E. C. H., Wong, M. N., Velleman, S. L., Lai, E. M. C., & Cheung, T. Y. K. (2023). *Structural and functional oral-motor abilities in typically developing Cantonese-speaking preschool children and children with childhood apraxia of speech*. The 19th Biennial Conference of the International Clinical Phonetics and Linguistics Association, Salzburg, Austria.

Wong, E. C. H., Wong, M. N., Velleman, S. L., Tong, M. C. F., & Lee, K. Y. S. (2023). Lexical tone perception and production in Cantonese-speaking children with childhood apraxia of speech: A pilot study. *Clinical Linguistics & Phonetics, 37*(4–6), 316–329. https://doi.org/10.1080/02699206.2022.2074310

Xu, Y. (2013). *ProsodyPro—A tool for large-scale systematic prosody analysis*. Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France.

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

Zhang, J. (2018). *A comparison of tone normalization methods for language variation research*. 32nd Pacific Asia Conference on Language, Information and Computation, Hong Kong.

Zhao, Y., & Jurafsky, D. (2009). The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics, 37*(2), 231–247. https://doi.org/10.1016/j.wocn.2009.03.002

Appendix A (p. 1 of 3)

Assessment Results of Participants in the Control Groups

ID	Group	Dx	Age	Sex	TOPOL (percentile)				CRVT	HKCAT (percentile)			
					SeC	SP	StC[a]	VN		IC	V/D	FC	T
B001	SSD	SSD	4;9	M	95	63	91	84	68.5	37	63	50	50
B002	SSD	SSD	5;3	F	91	25	63	63	62.0	63	63	2	50
B003	SSD	SSD	4;7	M	75	37	25	25	78.1	63	16	75	50
B004	SSD	SSD	3;1	F	50	98	NA	63	75.7	2	37	91	50
B005	SSD	SSD	4;4	F	95	95	91	84	97.3	63	16	75	50
B006	SSD	SSD	5;4	F	91	75	63	84	46.2	5	63	75	50
B007	SSD	SSD	4;9	M	75	16	16	63	51.7	50	16	50	50
B008	SSD	SSD	5;2	M	91	75	75	91	76.0	2	63	25	50
B009	SSD	SSD	3;9	F	50	63	NA	95	70.1	16	75	84	50
B011	SSD	SSD	4;7	F	75	84	75	84	51.7	63	16	50	50
B012	SSD	SSD	5;4	F	75	63	91	50	81.8	16	63	50	50
B013	SSD	SSD	3;6	M	75	98	NA	63	64.8	37	75	16	50
B014	SSD	SSD	4;8	F	95	25	95	91	68.5	75	1	50	50
B015	SSD	SSD	4;7	M	91	50	37	16	45.9	37	16	75	0.1
B016	SSD	SSD	3;4	F	98	75	NA	75	90.9	63	0.1	0.1	0.1
B017	SSD	SSD	5;10	M	25	37	75	37	32.8	9	63	63	50
B018	SSD	SSD	4;11	F	63	84	75	75	85.7	37	1	50	50
B020	SSD	SSD	5;6	M	84	91	95	25	53.0	37	5	9	50
B021	SSD	SSD	4;10	F	91	75	75	75	82.2	50	16	25	50
B022	SSD	SSD	4;10	F	84	16	91	37	57.5	63	63	0.1	50
B023	SSD	SSD	4;9	M	75	37	37	84	68.5	0.1	63	25	50
B024	SSD	SSD	3;1	F	99.6	63	NA	63	98.2	63	37	16	50
C001	S&LD	S&LD	3;10	M	50	2	NA	5	24.0	1	50	16	50
C002	S&LD	S&LD	4;6	M	37	9	37	5	51.7	9	16	75	50
C004	S&LD	S&LD	5;9	M	75	25	75	63	1.2	9	63	63	50
C007	S&LD	S&LD	5;5	M	37	5	2	0.1	1.8	0.1	0.1	25	50
C008	S&LD	S&LD	3;11	M	50	5	2	0.1	52.6	0.4	50	37	50
C010	S&LD	S&LD	4;8	F	37	16	9	37	63.1	1	0.1	9	0.1
A001	TD	WNL	3;6	M	91	75	NA	37	58.9	63	63	84	50
A002	TD	WNL	3;8	F	99.9	99.9	NA	84	94.7	63	75	63	50
A003	TD	WNL	3;7	F	99.6	98	NA	75	75.4	75	75	91	50
A004	TD	WNL	3;9	M	98	99	NA	63	90.3	63	75	91	50
A005	TD	WNL	3;0	M	99.9	99.9	NA	98	99.8	25	75	91	50
A006	TD	WNL	4;3	M	75	75	NA	95	86.6	75	75	91	50
A007	TD	WNL	4;8	M	98	84	91	91	63.1	63	63	91	50
A008	TD	WNL	4;4	F	95	84	95	95	92.5	75	63	91	50
A009	TD	WNL	5;2	F	98	63	25	25	81.8	63	63	50	50
A010	TD	WNL	5;6	F	84	84	91	95	72.5	63	63	91	50
A011	TD	WNL	5;1	M	95	95	84	84	76.0	75	75	91	50
A012	TD	WNL	5;8	M	91	84	84	37	72.5	63	63	91	50
A013	TD	WNL	5;1	M	63	37	25	37	54.2	25	63	91	50
A014	TD	WNL	3;10	M	95	75	NA	84	70.1	63	63	91	50
A016	TD	WNL	4;10	F	91	98	84	63	73.5	63	63	75	50
A017	TD	WNL	5;6	F	91	98	95	98	80.4	63	63	75	50

(*table continues*)

Appendix A (p. 2 of 3)

Assessment Results of Participants in the Control Groups

ID	Group	Dx	Age	Sex	TOPOL (percentile)				CRVT	HKCAT (percentile)			
					SeC	SP	StC[a]	VN		IC	V/D	FC	T
A018	TD	WNL	4;0	F	91	91	75	75	89.8	63	63	91	50
A019	TD	WNL	4;1	M	63	91	50	99	54.7	25	63	91	50
A020	TD	WNL	3;11	F	75	98	84	63	64.6	63	50	63	50
A021	TD	WNL	4;0	F	84	16	50	16	48.2	63	63	63	50
A022	TD	WNL	5;2	M	84	91	75	98	76.0	63	63	50	50
A023	TD	WNL	4;0	M	50	75	84	91	73.0	63	63	91	50
A024	TD	WNL	3;8	F	95	98	NA	75	64.8	25	63	91	50
A025	TD	WNL	4;11	M	91	63	63	75	63.1	63	63	50	50
A026	TD	WNL	3;8	M	16	75	NA	63	75.4	37	75	63	50
A027	TD	WNL	4;5	M	63	25	50	16	82.7	50	63	75	50
A028	TD	WNL	3;4	F	95	98	NA	95	95.7	75	63	91	50
A029	TD	WNL	4;2	F	99.6	95	37	16	78.2	63	50	63	50
A030	TD	WNL	4;2	F	99	95	98	99	78.2	50	63	37	50
A031	TD	WNL	3;8	F	63	75	NA	91	70.3	37	75	50	50
A032	TD	WNL	3;11	F	99	95	NA	84	83.9	50	37	84	50
A033	TD	WNL	4;1	M	84	91	98	75	96.1	63	63	91	50
A034	TD	WNL	4;5	M	99.6	91	95	37	92.5	63	63	50	50
A035	TD	WNL	4;4	M	99.6	50	37	91	67.3	50	63	50	50
A036	TD	WNL	4;3	M	91	75	75	64	78.2	75	63	63	50
A037	TD	WNL	4;5	F	99	99.9	98	91	86.6	50	63	50	50
A038	TD	WNL	5;8	F	91	50	98	95	80.4	63	63	37	50
A039	TD	WNL	3;4	F	99.9	91	NA	91	99.5	63	75	63	50
A040	TD	WNL	5;5	F	50	84	91	91	81.8	50	63	50	50
A042	TD	WNL	5;10	F	50	91	84	84	53.0	37	63	25	50
A043	TD	WNL	4;6	M	98	75	84	99	91.3	25	63	50	50
A044	TD	WNL	5;0	M	75	63	84	84	46.2	25	63	50	50
A045	TD	WNL	4;7	F	99.6	98	95	95	73.5	50	63	75	50
A046	TD	WNL	4;9	M	25	75	75	50	45.9	63	16	91	50
A047	TD	WNL	3;4	F	91	95	NA	25	67.4	75	25	50	50
A048	TD	WNL	3;3	M	98	84	NA	98	96.7	75	75	50	50
A049	TD	WNL	3;9	F	84	50	NA	37	70.1	75	75	50	50
A050	TD	WNL	3;3	F	99.6	91	NA	95	96.4	75	63	25	50
A052	TD	WNL	5;2	F	84	84	84	75	38.4	50	63	50	50
A053	TD	WNL	4;11	F	75	91	91	84	78.1	50	63	75	50
A054	TD	WNL	4;10	M	95	84	75	98	82.2	37	63	75	50
A055	TD	WNL	5;4	M	84	91	63	63	62.0	63	63	25	50
A056	TD	WNL	3;6	F	95	63	NA	95	52.8	75	75	91	50
A057	TD	WNL	3;4	M	84	50	NA	84	90.9	75	63	63	50
A058	TD	WNL	4;9	F	84	84	50	84	68.5	50	63	75	50
A059	TD	WNL	4;6	M	99	99	75	84	73.5	63	63	50	50
A060	TD	WNL	5;9	M	91	95	91	63	32.8	63	63	63	50
A061	TD	WNL	5;10	F	95	99	99	84	80.4	37	63	63	50
A062	TD	WNL	3;5	M	99.9	99.9	NA	99.6	98.7	25	75	63	50

(*table continues*)

Appendix A (p. 3 of 3)

Assessment Results of Participants in the Control Groups

ID	Group	Dx	Age	Sex	TOPOL (percentile)				CRVT	HKCAT (percentile)			
					SeC	SP	StC[a]	VN		IC	V/D	FC	T
A063	TD	WNL	3;3	M	75	98	NA	95	92.8	84	75	63	50
A064	TD	WNL	3;1	M	25	63	NA	63	48.5	63	37	84	50
A066	TD	WNL	3;8	F	37	16	NA	75	58.9	63	75	63	50
A067	TD	WNL	5;4	F	91	84	91	99	54.2	63	63	91	50

Note. Dx = diagnosis; TOPOL = Hong Kong Test of Preschool Oral Language (Child Assessment Service, 2020); CRVT = Hong Kong Cantonese Receptive Vocabulary Task (Lee et al., 2009); HKCAT = Hong Kong Cantonese Articulation Test (Cheung et al., 2006); SeC = Sentence Comprehension subtest; SP = Sentence Production subtest; StC = Story Comprehension subtest; VN = Vocabulary Naming subtest; IC = initial consonant; V/D = vowel or diphthong; FC = final consonant; T = tone; NA = not available; SSD = non-CAS speech sound disorder only; M = male; F = female; S&LD = non-CAS speech sound disorder plus language disorder; TD = children with typical speech-language development; WNL = within normal limits.

[a]The Story Comprehension subtest (StC) of the TOPOL was administered to children aged between 4;0 and 5;11 (years; months), as recommended in the test manual.

Appendix B

Results of Group Comparisons on Percentiles of Speech and Language Standardized Tests

Test/subtest	Kruskal–Wallis test		Post hoc pairwise comparisons[a]					
	df	*H*	CAS vs. S&LD	CAS vs. SSD	CAS vs. TD	S&LD vs. SSD	S&LD vs. TD	SSD vs. TD
TOPOL								
SeC	3	21.729*	−0.072	−2.352	−3.451*	−2.261	−3.353*	−1.579
SP	3	34.974*	−0.110	−2.340	−4.035*	−2.372	−4.232*	−2.889*
StC[b]	3	12.956*	0.352	−2.129	−2.995*	−1.525	−2.259	−1.165
VN	3	28.840*	−0.088	−2.528	−3.942*	−2.418	−3.823*	−2.098
CRVT	3	18.935*	0.426	−2.505	−3.463*	−1.972	−2.888*	−1.315
HKCAT								
IC	3	36.637*	−0.227	−2.507	−4.221*	−2.399	−4.269*	−2.904*
V/D	3	38.407*	−0.952	−2.338	−4.698*	−1.144	−3.411*	−3.757*
FC	3	30.283*	−1.450	−2.263	−4.423*	−0.445	−2.463	−3.421*
T	3	64.825*	−4.949*	−6.768*	−8.026*	−0.564	−1.338	−1.259

Note. CAS = childhood apraxia of speech; S&LD = children with non-CAS speech sound disorder plus language disorder; SSD = children with non-CAS speech sound disorder only; TD = children with typical speech-language development; TOPOL = Hong Kong Test of Preschool Oral Language (Child Assessment Service, 2020); SeC = Sentence Comprehension subtest; StC = Story Comprehension subtest; SP = Sentence Production subtest; VN = Vocabulary Naming subtest; CRVT = Hong Kong Cantonese Receptive Vocabulary Task (Lee et al., 2009); HKCAT = Hong Kong Cantonese Articulation Test (Cheung et al., 2006); IC = initial consonant; V/D = vowel or diphthong; FC = final consonant; T = tone.

[a]Bonferroni correction was applied. [b]The Story Comprehension subtest (StC) of the TOPOL was administered to children aged between 4;0 and 5;11 (years;months) as recommended in the test manual.

*$p < .05$.

Appendix C (p. 1 of 2)

Items From the Tone Sequencing Task

Table C1. Items included in the first set of the tone sequencing task.

Variable	Word stimulus	Meaning	Transcription	Paired nonword stimulus with transcription
V structure				
Tone 1	呀	Particle for exclamation	[aː1]	[ɛː1]
Tone 2	啞	Dumb	[aː2]	[ɛː2]
Tone 4	牙	Tooth/teeth	[aː4]	[ɛː4]
Tone 1	屙	Excretion	[ɔː1]	[iː1]
Tone 2	鵝	Goose	[ɔː2]	[iː2]
Tone 4	鵝	Goose	[ɔː4]	[iː4]
CV structure				
Tone 1	咩	Sound of sheep or question word	[mɛː1]	[jɛː1]
Tone 2	歪	Crooked	[mɛː2]	[jɛː2]
Tone 4	爺	Grandfather	[jɛː4]	[mɛː4]
Tone 1	多	More	[tɔː1]	[nɔː1]
Tone 2	朵	A classifier for flower	[tɔː2]	[nɔː2]
Tone 4	挪	take	[nɔː4]	[tɔː4]

Note. V = vowel; CV = consonant–vowel.

Table C2. Items included in the second set of the tone sequencing task.

Structure	Tone sequence[a]	Item		
Vowel (V)	124	[aː1ɔː2ɛː4]	[ɔː1ɛː2aː4]	[ɛː1aː2ɔː4]
	124	[aː1ɛː2ɔː4]	[ɔː1aː2ɛː4]	[ɛː1ɔː2aː4]
	142	[aː1ɔː4ɛː2]	[ɔː1ɛː4aː2]	[ɛː1aː4ɔː2]
	142	[aː1ɛː4ɔː2]	[ɔː1aː4ɛː2]	[ɛː1ɔː4aː2]
	214	[aː2ɔː1ɛː4]	[ɔː2ɛː1aː4]	[ɛː2aː1ɔː4]
	214	[aː2ɛː1ɔː4]	[ɔː2aː1ɛː4]	[ɛː2ɔː1aː4]
	241	[aː2ɔː4ɛː1]	[ɔː2ɛː4aː1]	[ɛː2aː4ɔː1]
	241	[aː2ɛː4ɔː1]	[ɔː2aː4ɛː1]	[ɛː2ɔː4aː1]
	412	[aː4ɔː1ɛː2]	[ɔː4ɛː1aː2]	[ɛː4aː1ɔː2]
	412	[aː4ɛː1ɔː2]	[ɔː4aː1ɛː2]	[ɛː4ɔː1aː2]
	421	[aː4ɔː2ɛː1]	[ɔː4ɛː2aː1]	[ɛː4aː2ɔː1]
	421	[aː4ɛː2ɔː1]	[ɔː4aː2ɛː1]	[ɛː4ɔː2aː1]
Consonant–vowel (CV)	124	[tɔː1mɛː2jɛː4]	[mɛː1jɛː2tːɔ4]	[jɛː1tɔː2mɛː4]
	124	[tɔː1jɛː2mɛː4]	[mɛː1tɔː2jɛː4]	[jɛː1mɛː2tɔː4]
	142	[tɔː1mɛː4jɛː2]	[mɛː1jɛː4tɔː2]	[jɛː1tɔː4mɛː2]
	142	[tɔː1jɛː4mɛː2]	[mɛː1tɔː4jɛː2]	[jɛː1mɛː4tɔː2]
	214	[tɔː2mɛː1jɛː4]	[mɛː2jɛː1tɔː4]	[jɛː2tɔː1mɛː4]
	214	[tɔː2jɛː1mɛː4]	[mɛː2tɔː1jɛː4]	[jɛː2mɛː1tɔː4]
	241	[tɔː2mɛː4jɛː1]	[mɛː2jɛː4tɔː1]	[jɛː2tɔː4mɛː1]
	241	[tɔː2jɛː4mɛː1]	[mɛː2tɔː4jɛː1]	[jɛː2mɛː4tɔː1]
	412	[tɔː4mɛː1jɛː2]	[mɛː4jɛː1tɔː2]	[jɛː4tɔː1mɛː2]

(*table continues*)

Table C2. *(continued).*

Structure	Tone sequence[a]	Item		
	412	[tɔː4jɛː1mɛː2]	[mɛː4tɔː1jɛː2]	[jɛː4mɛː1tɔː2]
	421	[tɔː4mɛː2jɛː1]	[mɛː4jɛː2tɔː1]	[jɛː4tɔː2mɛː1]
	421	[tɔː4jɛː2mɛː1]	[mɛː4tɔː2jɛː1]	[jɛː4mɛː2tɔː1]

[a]Tone sequence is presented in digits, which represent the Cantonese tones, that is, 1 is high-level tone, 2 is high-rising tone, and 4 is low-falling tone. The first digit denotes the initial tone, the second digit corresponds to the second tone, and the final digit signifies the last tone.

Table C3. Items included in the third set of the tone sequencing task.

Nonword stimulus	Transcription	Paired word stimulus	Meaning	Transcription
牙邊房	[aː4 piːn1 fɔːŋ2]	牙膏座	Toothpaste holder	[aː4 kou1 tsɔː2]
工子爐	[kʊŋ1 tsiː2 lou4]	哈比人	Hobbit	[haː1 peːi2 jɐn4]
打人汁	[taː2 jɐn4 tsɐp1]	搶銀包	To snatch wallets	[tsʰœːŋ2 ɐn4 paːu1]
提比包	[tʰɐi4 peːi2 paːu1]	提子汁	Grape juice	[tʰɐi4 tsiː2 tsɐp1]
哈銀座	[haː1 ɐn4 tsɔː2]	工人房	Worker's room	[kʊŋ1 jɐn4 fɔːŋ2]
搶膏人	[tsʰœːŋ2 kou1 jɐn4]	打邊爐	Eat hot pot	[taː2 piːn1 lou4]

Appendix D

R Codes Used in Analyses

Imputation for missing data

```
> library(mice)
> # imputation by using predictive mean matching
> imputed_DataX < – mice(all, method = ‘pmm’)
> completeDataX < – complete(imputed_DataX)
```

Analysis of *F*0 values of T1 syllables between the groups

```
> library(lme4)
> # create second-order orthogonal polynomial
> tl <– poly(unique(completeDataT1$Time), 2)
> # create orthogonal polynomial time variables in data frame
> completeDataT1[,paste(“ot”, 1:2, sep = ““)] < – tl[completeDataT1$Time, 1:2]
> # fit base model
> completeDataT1.base <– lmer(F0 ~ (ot1 + ot2) + (ot1 + ot2 | Subject), data=completeDataT1, REML = FALSE)
> # add effect of Group on intercept
> completeDataT1.0 < – lmer(F0 ~ (ot1 + ot2) + Group + (ot1 + ot2 | Subject), data=completeDataT1, REML = FALSE)
> # add effect of Group on linear term
> completeDataT1.1 < – lmer(F0 ~ (ot1 + ot2) + Group + ot1:Group + (ot1 + ot2 | Subject), data=completeDataT1, REML = FALSE)
> # add effect of Group on quadratic term
> completeDataT1.2 < – lmer(F0 ~ (ot1 + ot2)* Group + (ot1 + ot2 | Subject), data=completeDataT1, REML = FALSE)
> # compare models
> anova(completeDataT1.base, completeDataT1.0, completeDataT1.1, completeDataT1.2)
> # get parameter estimates and estimate p-values
> completeDataT1.coefs <– data.frame(coef(summary(completeDataT1.2)))
> completeDataT1.coefs$p < – 2*(1-pnorm(abs(completeDataT1.coefs$t.value)))
> completeDataT1.coefs
```

Pairwise comparisons of Tone 1 (high level) syllables among the groups

```
> # develop a contrast matrix for multiple comparisons
> contrast.matrix = rbind(“CAS vs. S&LD” = c(0,0,0,1,0,0,0,0,0,0,0,0), “CAS vs. SSD” = c(0,0,0,0,1,0,0,0,0,0,0,0), “CAS vs. TD” = c(0,0,0,0,0,1,0,0,0,0,0,0), “S&LD vs. SSD” = c(0,0,0,-1,1,0,0,0,0,0,0,0), “S&LD vs. TD” = c(0,0,0,-1,0,1,0,0,0,0,0,0), “SSD vs. TD” = c(0,0,0,0,-1,1,0,0,0,0,0,0), “Slope: CAS vs. S&LD” = c(0,0,0,0,0,0,1,0,0,0,0,0), “Slope: CAS vs. SSD” = c(0,0,0,0,0,0,0,1,0,0,0,0), “Slope: CAS vs. TD” = c(0,0,0,0,0,0,0,0,1,0,0,0), “Slope: S&LD vs. SSD” = c(0,0,0,0,0,0,-1,1,0,0,0,0), “Slope: S&LD vs. TD” = c(0,0,0,0,0,0,-1,0,1,0,0,0), “Slope: SSD vs. TD” = c(0,0,0,0,0,0,0,-1,1,0,0,0), “Degree: CAS vs. S&LD” = c(0,0,0,0,0,0,0,0,0,1,0,0), “Degree: CAS vs. SSD” = c(0,0,0,0,0,0,0,0,0,0,1,0), “Degree: CAS vs. TD” = c(0,0,0,0,0,0,0,0,0,0,0,1), “Degree: S&LD vs. SSD” = c(0,0,0,0,0,0,0,0,0,-1,1,0), “Degree: S&LD vs. TD” = c(0,0,0,0,0,0,0,0,0,-1,0,1), “Degree: SSD vs. TD” = c(0,0,0,0,0,0,0,0,0,0,-1,1))
> library(multcomp)
> compst1 < – glht(T1.0, contrast.matrix)
> summary(compst1)
```

Note. The bold codes indicate the specific parameter for the analysis.