ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Pitch Variation Skills in Cantonese Speakers With Apraxia of Speech After Stroke: Preliminary Findings of Acoustic Analyses

Eddy C. H. Wong,[a] Min Ney Wong,[a,b,c,d] Si Chen,[a,b,c,d] and Joyce Y. W. Lin[a]

[a] Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, China [b] Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, China [c] Research Institute for Smart Ageing, The Hong Kong Polytechnic University, China [d] The HK PolyU-PekingU Research Centre on Chinese Linguistics, The Hong Kong Polytechnic University, China

ARTICLE INFO

Article History:
Received April 12, 2023
Revision received July 3, 2023
Accepted September 9, 2023

Editor-in-Chief: Cara E. Stepp
Editor: Raymond D. Kent

https://doi.org/10.1044/2023_JSLHR-23-00242

ABSTRACT

Purpose: Literature on apraxia of speech (AOS) in Chinese speakers is sparse compared to the English literature. This study aims to examine the pitch variation skills of Cantonese adults with AOS poststroke in terms of perceptual tone accuracy, acoustic fundamental frequency (f_o) changes, and repetition durations on items with different syllable structures, lexical status, and tone syllables in various positions in a sequencing context.

Method: Six Cantonese adults with AOS poststroke (AOS group), six adults without AOS poststroke (nAOS group), and six healthy controls (HC group) performed the tone sequencing task (TST), which was adapted from oral diadochokinetic tasks, with three different tone syllables. Tone accuracy, f_o values across 10 time points, and acoustic repetition durations were compared within and between the groups.

Results: The AOS group produced significantly lower tone accuracy and different f_o changes on the three Cantonese tone syllables compared with the control groups and significantly longer repetition durations than the HC group. The AOS group showed more difficulty with the tone syllables with the consonant–vowel structure, while a priming effect was observed on the T2 (high-rising) syllables with lexical meanings. A unique lowering of f_o in the final syllable of the trisyllabic items was observed only in the AOS group.

Conclusions: The AOS group showed degraded pitch variation skills. The effects of the three linguistic elements were discussed. Future investigations are called for to adapt the TST in other tonal languages to determine if degraded pitch variation skills are present in other tonal language speakers with AOS.

Apraxia of speech (AOS), a type of acquired neurological motor speech disorder, can be found after a stroke and commonly co-occurs with aphasia and/or dysarthria (Duffy, 2020). AOS is characterized by frequent and irregular articulatory errors; speech behaviors, such as articulatory groping; and inappropriate prosody (Ballard et al., 2014; Duffy et al., 2020). Inappropriate prosody in English speakers with AOS is observed as slow speech rate, lengthened sound segments, syllable segmentation within and between words, equal stress, reduced pitch and loudness variability, abnormal intonation, and lexical stress errors (Duffy, 2020; McNeil et al., 2009; Strand et al., 2014; Utianski et al., 2018; Wambaugh et al., 2006). Lexical stress errors refer to the inappropriate assignments of stress in words. Lexical stress patterns are formed by the simultaneous variation of three parameters, namely, length, pitch, and loudness (the physical attributes are duration, fundamental frequency [f_o], and intensity, respectively). Acoustic measures provide quantitative data

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

about lexical stress to assist in the diagnoses of AOS in children (Terband et al., 2019). In adults, the pairwise variability index for vowel (V) duration and word syllable duration has been shown to be useful in differentiating adults with and without AOS poststroke (Haley & Jacks, 2019; Vergis et al., 2014). Studies that have examined individuals with AOS placed greater emphasis on duration measurements as opposed to pitch, since variations in f_o had been determined to be the least reliable indicator in differentiating English speakers with and without AOS (Ballard et al., 2012; Haley & Jacks, 2019; Kopera & Grigos, 2020; Vergis et al., 2014).

Unlike English, measuring lexical stress might not be useful in Cantonese speakers, in that pitch variational measures might provide more valuable information for identification and understanding of the clinical manifestations of motor speech disorders in Cantonese speakers. Cantonese is characterized by lexical tones, in which different pitch heights (i.e., high, mid, and low) and contours (i.e., level, rising, and falling) indicate different lexical meanings (Yip & Matthews, 2011). There are six contrastive tones in Cantonese, including high-level (T1), high-rising (T2), midlevel (T3), low-falling (T4), low-rising (T5), and low-level (T6) tones. These tones are associated with vowels. Despite the significance of tone/pitch in Cantonese, research on tone production in Cantonese speakers with motor speech disorders is scarce. In a perceptual study of Cantonese speakers with hypokinetic dysarthria, another type of acquired neurological motor speech disorder associated with Parkinson's disease, Whitehill et al. (2003) found that the participants' lexical tone productions were relatively unaffected. Recently, Tee et al. (2022) reported tonal substitution errors in a Cantonese–English bilingual woman with AOS associated with the nonfluent variant of primary progressive aphasia. In conjunction with neuroimaging results, the authors speculated that tone production impairment is the manifestation of impaired motor speech planning ability around the vocal cord area (Tee et al., 2022). A similar claim was proposed for Cantonese-speaking children with AOS (E. C. H. Wong et al., 2020, 2021, 2022). These studies provided a basis for further investigation of pitch variation skills in Cantonese speakers with motor speech disorders, especially AOS.

To the authors' knowledge, Tee et al. (2022) was the first ever investigation of tone production skills in Cantonese speakers with AOS. The study included a serial tone reading task, tone–word reading tasks, and consecutive tone repetition tasks. Acoustic analysis was used to gather data from the serial tone reading task, and spectrograms were visually inspected to compare f_o values and changes over time between the participant and controls. The results revealed that the participant substituted the low-falling tone (T4) with the high-rising tone (T2). However, the results were not verified by any statistical analyses, and the stimuli used were not controlled for by syllable structure, lexical status, or tone syllable position.

Syllable structure, lexical status, or tone syllable position could have significant impacts on pitch variation skills in Cantonese speakers with AOS. Cantonese syllable structure ranges from the less complex consonant (C) or V structure to the most complex CVC structure. Although the slope of f_o rising in Mandarin (another tonal language with four lexical tones, including high-level, rising, low, and falling tones) tones did not vary with syllable structure in healthy adults (Xu, 1998), the relatively complex syllable structures (i.e., CVC or CV) might bring extra challenges to the speaker with AOS. 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, and speakers with AOS have difficulty containing multiple phonological units at a time, specifically in the phonological content buffer (Guenther, 2016; Miller & Guenther, 2020). Speakers with AOS may have difficulty in managing both syllable structure and tone during speech production, resulting in maintaining syllable structure accuracy but sacrificing tone accuracy, or vice versa. Cantonese speakers with AOS might have more difficulty in managing pitch and relatively complex syllable structures during speech processing.

In addition, given that a speaker's and listener's lexical knowledge plays a role in lexical tone production and perception (Lee et al., 2002; Zhao & Jurafsky, 2009), the semantic priming of word items may enhance listeners' perception of a given word, even if that word was produced with degraded pitch variation. Using meaningful word items might result in overlooking actual degraded pitch variation skills in speakers with AOS. Investigation of pitch variations in Cantonese speakers with AOS should, therefore, involve nonword items. Inclusion of items with different lexical status in the investigation of pitch variation skills could further examine the presence of priming effects in the speech of speakers with AOS.

Finally, it might be important to consider tone syllable position when examining the pitch variation skills of Cantonese speakers with AOS. The pitch of a syllable can be influenced by the preceding and subsequent tone syllables through carryover and anticipatory tonal coarticulation (Li & Chen, 2016), in terms of pitch height (f_o values), slope (rate of change of f_o values), and curvature (rate of change of pitch slope). When evaluating pitch variation skills, it is important to control the adjacent tone syllables, which can be achieved by balancing the tone syllables in different positions. Moreover, research has indicated that the pitch contours of tones may vary depending on their position within a sentence. For instance, studies on Taiwan Mandarin have shown that the rising tone in sentence-final position has a higher pitch height at the offset compared to the onset; however, in sentence-medial position, the rising tone is more like a falling–rising tone (i.e., the onset is higher than the offset, and the falling portion is longer than the rising portion; Fon & Hsu, 2004). Furthermore, in sentence-final position, the rising tone has a lower pitch height and steeper pitch slope than that in sentence-medial position (Fon & Hsu, 2004). Therefore, when examining the pitch variation skills of Cantonese speakers with AOS, it is important to consider the effects of syllable position.

E. C. H. Wong et al. (2021) examined the effects of syllable structure and lexical status using the tone sequencing task (TST), which is a pitch-sequencing task adapted from the oral diadochokinetic (DDK) tasks. In contrast to the oral DDK tasks, where speakers are required to rapidly repeat one and three CV syllables with different Cs, the TST involves rapid repetition of one- and three-tone syllables in sequence. The outcome measures include f_o and repetition duration, and perceptually judged tone accuracy. E. C. H. Wong et al. investigated the pitch variation skills of three children with AOS and six children without AOS using the TST. The results showed that children with AOS produced significantly longer repetition durations on the TST items, but there were no between-groups differences on items with different syllable structures (V vs. CV) nor lexical status (words vs. nonwords). All of the groups used similar lengths of time to complete word and nonword items and a longer time to complete items with V than CV structures (E. C. H. Wong et al., 2021). It was explained that the separation of V syllables in the tone repetition task is achieved by cessation of phonation resulting from vocal fold closure, while the separation of CV syllables is achieved using oral articulators to construct consonantal constriction in the vocal tract. The authors suggested that acoustic repetition duration is useful in revealing pitch variation difficulty in children with AOS. The study also provided a method to investigate pitch variation skills and the effects of linguistic variables on pitch variation in Cantonese speakers with AOS.

The purpose of this study is to examine pitch variation skills in adults with AOS poststroke using perceptual judgment of tone accuracy, acoustic analyses of f_o values, and repetition duration in the context of the TST. Four hypotheses were developed: (a) The f_o values of tone syllables produce by the adults with AOS poststroke (AOS group) would differ from those produced by the control groups in terms of pitch height, slope, and curvature. In particular, the AOS group would show a flatter pitch change across time in the contour tones (high-rising [T2] and low-falling [T4]) when compared with the control groups. A falling contour in T4 (low-falling) syllables would be substituted by T2 (high-rising) syllables, as observed in the study of Tee et al. (2022). In addition, the AOS group would have lower tone accuracy and longer repetition duration compared with the control groups. (b) The AOS group would produce different f_o values on items with different syllable structures, while the control groups would not. In particular, the control groups would produce similar f_o values across time on items with V and CV structures, given that pitch did not vary with syllable structures in Mandarin healthy speakers (Xu, 1998). The AOS group would produce a flatter pitch change across time and lower tone accuracy on items with CV structure than items with V structure, given that they have difficulty in simultaneous management of syllable structure and tones. Moreover, the groups would produce longer repetition duration on items with V structure than those with CV structure, as reported in Cantonese-speaking children with AOS (E. C. H. Wong et al., 2021). (c) There would be differences in pitch performance between items with different lexical status within each group. In particular, the lexical status would facilitate the production of f_o values and tone accuracy, but not repetition duration. (d) All of the groups would produce different f_o values on syllables in different positions (i.e., initial vs. medial vs. final), given that healthy speakers produce pitch differently in different positions. In addition, it is anticipated that there would be a greater number of pitch performance differences (i.e., in pitch heights, slopes, and curvatures) among the syllables in different positions produced by the AOS group compared with those produced by the control groups.

Method

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

Participants

Six adults with AOS poststroke were recruited (AOS group). The inclusion criteria of participants included (a) aged 18 years or above, (b) speaking Cantonese as the dominant language for communication, (c) having no uncorrected visual impairment, (d) having at least one ear that passed the hearing screening, (e) having no structural abnormality that affects speech, and (f) a confirmed AOS diagnosis by the two speech-language pathologists (SLPs) described in the next section. The exclusion criteria included (a) having no AOS after stroke, (b) failing the hearing screening in both ears, (c) having a limited verbal output, and (d) being unable to complete all of the assessment tasks. Six adults without AOS poststroke but with aphasia with or without co-existing dysarthria (nAOS group) and six healthy controls (HC group) were also recruited. The controls were age-matched (± 6 years) with the participants in the AOS group. The demographic information of the participants is presented in Table 1. There was no significant difference between the groups in terms of age, $H(2) = 0.978$, $p = .613$, and no significant difference between the AOS and nAOS groups in terms of the number of months postonset ($U = 1.087$, $p = .297$). All of the participants provided written consent.

Confirmation of Diagnoses

All of the participants received a 2-hr speech and language assessment by one of the two master speech-language pathology students under the supervision of a qualified SLP (first author). The assessment protocol included case history taking, conversational speech sample collection, the Cantonese Aphasia Battery (CAB; Yiu, 1992), the Computerized Revised Token Test–Cantonese (CRTT-Cantonese; Bakhtiar et al., 2020), the unofficial Cantonese adapted version of the Frenchay Dysarthria Assessment–Second Edition (FDA-2; Enderby & Palmer, 2008), the Cantonese Tone Identification Test (Lee, 2012), a set of informal speech production tasks (including a repeated trials task, a word and nonword repetition task, and an increased word length task), and DDK tasks, which included alternating motor rates and sequential motion rates (Pierce et al., 2013). The sessions were video-recorded. The diagnosis of aphasia was confirmed if the participant fell below the cutoff point for the CAB Aphasia Quotient (< 96.4) and the cutoff points for the Cantonese Reading Comprehension version (CRTT-R-$_{WF}$-Cantonese; < 13.12) and the Cantonese Listening Comprehension version (CRTT-L-Cantonese; < 13.75) of the CRTT. There were no significant differences between the AOS and nAOS groups on the CAB Aphasia Quotient ($U = 23.00$, $p = .485$), the CRTT-L-Cantonese scores ($U = 6.00$, $p = .700$), or the CRTT-R-$_{WF}$-Cantonese scores ($U = 7.00$, $p = .400$).

The AOS diagnoses of the participants were confirmed with the consensus of two independent SLPs, who were blinded to the purpose of this study and the speech and language diagnoses of the participants, with the aid of the Apraxia of Speech Rating Scale 3.0 (Utianski et al., 2018). After the initial independent ratings, an online postrating meeting was held to produce a single consensus rating to be used in all further analyses. The assessment results of each participant are presented in Table 1, and the characteristics of each participant with AOS are presented in Appendix A.

The dysarthria diagnoses were given based on performance on the unofficial adapted version of FDA-2, as judged by an SLP (the first author). The cranial nerve (CN) lesions, dysarthria types, and severity levels of the participants are presented in Table 1. Three out of six participants with AOS poststroke (Participants 1, 2, and 5) had dysarthria with lesions on CN V (trigeminal nerve), VII (facial nerve), and XII (hypoglossal nerve). Participant 1 had upper motor neuron dysarthria, Participant 2 had flaccid and hypokinetic dysarthria, and Participant 5 had flaccid dysarthria. Two of these participants (1 and 2) had an additional lesion on CN X (vagus nerve). As this study focuses on examining pitch variation skills, information about the participants' laryngeal function is reported. The results of the FDA-2 laryngeal subtest showed that Participant 1 had difficulty with sustained phonation (rated as "c" among the five descriptors ranging from "a" [normal function] to "e" [no function]), loudness variation (rated as "d"), and voice production (rated as "b"), with no abnormality reported in pitch control. Participant 2 had difficulty with sustained phonation (rated as "b"), loudness variation (rated as "d"), and voice production (rated as "c"), with no abnormality reported in pitch control. Thus, despite experiencing difficulties with some laryngeal functions, neither Participant 1 nor Participant 2 exhibited any abnormalities in pitch control. Therefore, we believe that their dysarthria did not impact the results.

Procedure

Each participant underwent a data collection session in a sound-proof booth following their initial assessment on a single day. During this session, they were asked to perform four speech tasks: picture-naming and repeated trial task, word and nonword repetition task, increased word length task, and the TST. For this study, only the TST was analyzed. All tasks were conducted in a randomized sequence. The TST comprised three sets of items that 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 was given for each item. All models of the item and sequence were prerecorded by a healthy Cantonese-speaking male and presented to the participants auditorily via PowerPoint. The productions were audio-recorded using Praat software (Boersma & Weenink, 2022) with Audio-Technica (model: AT2035) or Shure (model: SM48) microphones and Steinberg (model: UR22mkll) or M-Audio (model: M-Track Plus II) interfaces.

TST

The TST used in this study was the third version, which was modified based on previous versions (E. C. H.

Table 1. Demographic information of the participants.

Group	ID	Gender	Age	Months postonset	CAB AQ	Aphasia type	CRTT-L-Cantonese score	CRTT-R-$_{WF}$-Cantonese score
AOS	#1	M	52	94	76	Conduction	NA	NA
	#2[a]	M	50	84	96.2	Anomic	NA	NA
	#5	M	56	165	82.3	Anomic	10.36	10.76
	#10	M	43	61	47	Broca's	10.55	7.48
	#12	M	62	21	12.9	Global	NA	NA
	#17	F	58	19	80.8	Anomic	11.90	12.74
M (*SD*)	NA	NA	53.50 (6.69)	74 (54.38)	65.87 (30.59)	NA	10.94 (0.84)	10.33 (2.66)
nAOS	#3	M	53	31	76.2	Transcortical motor	NA	NA
	#4	F	55	63	72.6	Transcortical sensory	9.88	9.39
	#7	F	49	58	95.8	Anomic	12.36	12.80
	#9	M	59	52	91.2	Anomic	12.74	12.94
	#11	M	64	19	77	Anomic	NA	NA
	#18	M	62	39	96	Anomic	NA	NA
M (*SD*)	NA	NA	57.00 (5.69)	43.67 (16.97)	84.80 (10.69)	NA	11.66 (1.55)	11.71 (2.01)
HC	#13	M	59	NA	99	NA	NA	NA
	#14	M	64	NA	98.4	NA	NA	NA
	#15	F	54	NA	100	NA	NA	NA
	#19	M	45	NA	100	NA	14.5	14.65
	#20	M	50	NA	99.8	NA	14.0	14.29
	#21	M	51	NA	100	NA	14.2	14.35
M (*SD*)	NA	NA	56.83 (6.79)	NA	99.3 (0.68)	NA	14.23 (0.25)	14.43 (0.19)
Group	**ID**	**Dysarthria type**		**CN lesions**		**Dysarthria severity**	**Apraxia of speech severity**	**CANTIT percentile (%)**
AOS	#1	UUMN		VII, X, XII		Mild	Mild	< 4
	#2[a]	Flaccid & hypokinetic		V, VII, X, XII		Mild	Mild	16
	#5	Flaccid		V, VII, XII		Mild	Mild	16
	#10	NA		NA		NA	Moderate	< 4
	#12	NA		NA		NA	Severe	< 4
	#17	NA		NA		NA	Mild	< 4
M (*SD*)	NA	NA		NA		NA	NA	NA
nAOS	#3	Flaccid		V, VII, X, XII		Mild	NA	44
	#4	Flaccid		V, VII, X, XII		Mild	NA	16
	#7	NA		NA		NA	NA	44
	#9	NA		NA		NA	NA	82
	#11	NA		NA		NA	NA	44
	#18	UUMN		Nil		Mild	NA	82
M (*SD*)	NA	NA		NA		NA	NA	52 (25.64)
HC	#13	NA		NA		NA	NA	44
	#14	NA		NA		NA	NA	82
	#15	NA		NA		NA	NA	44
	#19	NA		NA		NA	NA	82
	#20	NA		NA		NA	NA	44
	#21	NA		NA		NA	NA	82
M (*SD*)	NA	NA		NA		NA	NA	63.00 (20.81)

Note. CAB = Cantonese Aphasia Battery; AQ = Aphasia Quotient; CRTT-L-Cantonese = Computerized Revised Token Test-Listening Comprehension Version–Cantonese; CRTT-R-$_{WF}$-Cantonese = Computerized Revised Token Test-Reading-Word Fade Comprehension Version–Cantonese; AOS = adult speakers with apraxia of speech poststroke; NA = not applicable; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls; CN = cranial nerve; CANTIT = Cantonese Tone Identification Test (Lee, 2012); UUMN = unilateral upper motor neuron dysarthria; Flaccid = flaccid dysarthria; Hypokinetic = hypokinetic dysarthria; V = trigeminal nerve; VII = facial nerve; X = vagus nerve; XII = hypoglossal nerve.

[a]Participant 2 had encephalitis 1 month after the stroke.

Wong et al., 2019, 2021). There were 108 items, which were formed with early-acquired initial Cs (i.e., [m, j, t, n]) and Vs (i.e., [a, ɔ, i, ɛ]), as well as three early-acquired Cantonese tones, T1, T2, and T4. These three tones were chosen because they involve the upper and lower limits (i.e., the high and low tones) and all three tone contours in Cantonese (i.e., level, rising, and falling). The TST comprised three sets, including (a) V and CV monosyllabic words and nonwords, (b) V and CV trisyllabic nonwords, and (c) trisyllabic words and nonwords. The first set included 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 were within the top 21% frequency rank of Hong Kong Cantonese–speaking adults' speech, with a mean and standard deviation of 6.2% and 7.7%, respectively (Lai & Winterstein, 2020), except for three words, /ɔ1/ (means excretion), /ɔ2/ (means goose), and /jɛ4/ (means grandfather). However, these three words were found to be within the top 2.3% (M = 1.2%, SD = 1.0%) frequency rank of Hong Kong Cantonese–speaking children's speech (Lai & Winterstein, 2020). The items are presented in Appendix B, Table B1. The second set included 72 items, which were derived from three V structures (i.e., [a], [ɔ], and [ɛ]), three CV structures (i.e., [mɛ], [jɛ], and [tɔ]), and three Cantonese tones (i.e., T1, T2, and 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. The Cs and Vs used in this set were within the top 50% frequency rank among the 19 Cs and eight monophthongs of the language, while the T1, T2, and T4 tones fall at the first, fourth, and fifth/sixth frequency rank in three Cantonese corpora, respectively (Andrus et al., 2016; Leung & Law, 2001; Luke & Wong, 2015). All of these items were nonwords in Cantonese. The items are presented in Appendix B, Table B2. The third set 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 15% (M = 5.2%, SD = 4.6%) of the frequency rank in adults' speech (Lai & Winterstein, 2020), except for two disyllables (/a4 kou1/ ["toothpaste"] and /tʰɐi4 tsi2/ ["grape"], which, however, were found to be within the top 0.7% (M = 0.64%, SD = 0.03%) frequency rank of children's speech (Lai & Winterstein, 2020). The nonwords were formed using the syllables included in the words. Again, each item included all three tones, with the order of the tones varying from one item to the next. More details about these items are provided in Appendix B, Table B3. The difference between the second and third sets was the syllable structure. Only the 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.

Data Analyses

Three native Cantonese-speaking final-year master's students in speech-language pathology perceptually judged the performance of the participants using a binary system (i.e., 1 = *correct*, 0 = *incorrect*). The raters were blinded to the purpose of this study and the speech and language diagnoses of the participants. There were a total of 1,380 judgments for each participant on the TST. The items were rated as "correct" when two or more raters rated them as correct; otherwise, the items were rated as "incorrect."

Acoustic analyses were performed for the first completed set of five repetitions of all of the items from all of the participants to measure f_o values and repetition duration using Praat software (Boersma & Weenink, 2022). The audio segments of the videos were imported into Praat. Following the procedure suggested by Ma et al. (2006), the vocalic segment of each syllable was identified manually from a wideband spectrogram, and an amplitude waveform was visually displayed in the software. Eleven evenly spaced time points were identified from the beginning to the end of the vocalic segment of each syllable using an autocorrelation algorithm (0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of the total duration of each syllable) using ProsodyPro (Xu, 2013). No manual corrections were implemented to extract f_o contours. The f_o values at the beginning (i.e., 0%) were excluded from the analyses. Normalization of the observed f_o was conducted using the z-score method (Rose, 1987; Zhang, 2018). The normalized f_o values were calculated per participant by dividing the difference between observed f_o values and the mean by the standard deviation. Repetition duration was measured as the difference in time between the 0% time point of the first syllable (beginning of the set) and the 100% time point of the final syllable (end of the set) by using another Praat script (Lennes, 2002).

Missing values were observed due to technical issues during data collection, incomplete productions of five repetitions, or overlapping productions between the participants and models/assessors. In cases where participants failed to produce five repetitions, the produced repetitions were analyzed, and the rest were treated as missing values. Overlapping productions were not analyzed and were considered as missing values. To address randomly missing f_o values, predictive mean matching, which calculates the predicted value of missing normalized f_o according to the specific imputation model, was performed. The method formed a small set of candidate donors from all complete data sets that have predicted values closest to the predicted value for the missing entry. One donor was randomly drawn from the candidates, and the observed value of the donor was taken to replace the missing value. The imputation was performed in R using the Multivariate Imputation by Chained Equations (mice) package (Version 3.14.0). There were 1.64% missing values among the 248,400 expected normalized f_o values. Tables 2 and 3 list the percentage of missing data handled by predictive mean matching in each analysis and for each participant, respectively.

Effects of Group

Growth curve analysis (Mirman, 2014) was conducted to analyze the time course of normalized f_o values for each of the three Cantonese tones (i.e., T1, T2, and T4) among the groups. All of the productions from all three sets of the TST were included in the analyses. The overall time course of normalized f_o values was captured with a second-order orthogonal polynomial with fixed effects of group (AOS vs. nAOS vs. HC) on the 10 time points for each tone (from 10% to 100%). A total of three models were formed. The AOS group was treated as the baseline, and parameters were estimated for the nAOS and HC groups. Each model included random effects of participants on all of the time terms. The base of the model is presented with the function, F0 ~ (ot1 + ot2) + (ot1 + ot2 | Subject), data = Tone, REML = FALSE. The fixed effects of groups on intercept, a linear term, and a quadratic term were added individually. The functions are presented as follows:

Intercept: F0 ~ (ot1 + ot2) + Group + (ot1 + ot2 | Subject), data = Tone, REML = FALSE

Linear term: F0 ~ (ot1 + ot2) + Group + ot1: Group + (ot1 + ot2 | Subject), data = Tone, REML = FALSE

Quadratic term: F0 ~ (ot1 + ot2) × Group + (ot1 + ot2 | Subject), data = Tone, REML = FALSE

The effects on model fit were evaluated using model comparisons. Parameter-specific *p* values were estimated using the normal approximation (i.e., treating the *t* values as a *z* value). Multiple comparisons for the effects of syllable position (i.e., initial vs. medial vs. final position) were performed using the multcomp package (Version 1.4-20). All analyses were carried out in R (Version 4.2.0) using the lme4 package (Version 1.1-29). The Kruskal–Wallis test was used to compare the means of perceptual judgments and acoustic durations among the three groups. Mann–Whitney *U* tests with the Bonferroni correction were applied for post hoc analyses. The R code of the analysis of normalized f_o values for T1 between the groups is shown in Appendix C.

Table 2. Percentage of missing data handled by predictive mean matching for each analysis.

Variable	Expected number of data	Missing data (%)
Effects of group		
Tone 1	82,800	1.62
Tone 2	82,800	1.54
Tone 4	82,800	1.79
Effects of syllable structure		
AOS group	72,000	2.97
nAOS group	72,000	1.98
HC group	72,000	0.04
Effects of lexical status		
AOS group	82,800	3.01
nAOS group	82,800	1.90
HC group	82,800	0.04
Effects of syllable position		
AOS group	75,600	3.24
nAOS group	75,600	2.02
HC group	75,600	0.04

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls.

Effects of Linguistic Elements

The effects of parameters on normalized f_o values within each group were examined using growth curve analysis (Mirman, 2014). The parameters included syllable structure, lexical status, and tone syllable position. The effects of syllable structure were examined by comparing performance between items with V and CV structures from the first and second sets. The effects of lexical status were examined by comparing performance between word and nonword items from the first and third sets. The effects of syllable position were examined by comparing performance among the initial, medial, and final positions using all three sets. The overall time course of normalized f_o values was captured with a second-order orthogonal polynomial with fixed effects of the parameters (i.e., V vs. CV structure, word vs. nonwords, and initial vs. medial vs. final position). The fixed effects of the parameters (i.e., lexical status, syllable structure, or syllable position) on the intercept, linear term, and quadratic terms were added individually, and their effects on model fit were evaluated using model comparisons. The functions were revised according to the parameter to perform the appropriate analyses. Parameter-specific *p* values were estimated using the normal approximation (i.e., treating the *t* values as a *z* value). All analyses were carried out in R Version 4.2.0 using the lme4 package (Version 1.1-29). Multiple comparisons for the effects of syllable position (i.e., initial vs. medial vs. final position) were performed using the multcomp package (Version 1.4-20). The R codes of multiple comparisons of tone syllables in different positions are presented in Appendix C. Effects of syllable structure and

Table 3. Percentage of missing data handled by predictive mean matching for each participant.

ID	Group	Number of missing values out of 13,800 values per participant	Missing data (%)
#1	AOS	370	2.68
#2	AOS	90	0.65
#5	AOS	10	0.07
#10	AOS	200	1.45
#12	AOS	1,800	13.04
#17	AOS	20	0.14
#3	nAOS	50	0.36
#4	nAOS	240	1.74
#18	nAOS	20	0.14
#7	nAOS	190	0.38
#9	nAOS	960	6.96
#11	nAOS	70	0.51
#13	HC	30	0.22
#14	HC	0	0.00
#15	HC	0	0.00
#19	HC	0	0.00
#20	HC	0	0.00
#21	HC	0	0.00

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls.

lexical status on perceptual judgments and acoustic duration within each group were examined using the Mann–Whitney *U* tests.

Mixed-Effects Binomial Logistic Regression

Four logistic mixed models (estimated using ML and BOBYQA optimizer) were fitted to predict perceptual judgment of tone accuracy (Accuracy) with syllable structure (Structure: V structure vs. CV structure), lexical status (Lexical: word vs. nonword items), and syllable position (Position: initial vs. medial vs. final positions). The model included speakers as a random effect (formula: ~1 | Participant). Standardized parameters were obtained by fitting the model on a standardized version of the data set. Confidence intervals (CIs) and *p* values were computed using a Wald *z*-distribution approximation. The analyses were performed in R (Version 4.2.0). The R codes are presented in Appendix C.

Reliability Measures

The three speech-language pathology students rerated the tone accuracy of 11% of all of the judgments 1 week after the initial rating. The intrarater reliability of the three raters was calculated using two-way mixed effects, intraclass correlation coefficient (ICC; 3, 1). Good-to-excellent reliability (Koo & Li, 2016) was reported for the three raters: .88, 95% CI [0.75, 1.01], $p < .001$; .90, 95% CI [0.90, 0.90], $p < .001$; and .83, 95% CI [0.83, 0.84], $p < .001$.

Annotations were performed by a research assistant and an undergraduate student who were studying speech and language sciences. Both received a 2-hr training from a research assistant, who had extensive experience in the acoustic analysis of Cantonese speakers and obtained more than .80 on the ICC with the trainer on the f_o values of an adult after stroke (not included in this study) on the TST before their independent analyses. The ICC values obtained for training were .83 (95% CI [0.82, 0.83], $p < .001$) and .85 (95% CI [0.85, 0.86], $p < .001$), which demonstrated good reliability. Intra- and interrater reliabilities on annotation were calculated with approximately 10% of total re-annotations performed by the annotators using two-way random effects, ICC (2, 1). Excellent interrater (.91, 95% CI [0.91, 0.91], $p < .001$) and intrarater (.99, 95% CI [0.99, 0.99], $p < .001$, and .96, 95% CI [0.96, 0.96], $p < .001$) reliabilities were obtained.

Results

Effects of Group

Perceptual Judgment

Table 4 presents the inferential statistics of the perceptual judgments of tone accuracy in the three groups.

Table 4. Results of statistical tests for perceptual tone accuracy in adult speakers with apraxia of speech poststroke (AOS group), adult speakers without apraxia of speech poststroke (nAOS), and healthy controls (HC group).

Effects of group	Kruskal–Wallis test		Multiple comparison[a]		
	df	*H*	AOS vs. nAOS	AOS vs. HC	nAOS vs. HC
High-level (T1)	2	1,006.154***	−19.765***	−31.368***	11.603***
High-rising (T2)	2	3,154.645***	−38.859***	−54.550***	15.691***
Low-falling (T4)	2	1,877.723***	−29.312***	−42.295***	12.983***
Effects of syllable structure (V vs. CV structure)					
	Mann–Whitney *U* test				
	df	AOS	nAOS	HC	
High-level (T1)	1	−1.631	3.198**	1.000	
High-rising (T2)	1	−2.207*	−4.304***	−1.873	
Low-falling (T4)	1	−0.735	−0.113	0.000	
Effects of lexical status (word vs. nonwords)					
	Mann–Whitney *U* test				
	df	AOS	nAOS	HC	
High level (T1)	1	9.531***	5.103***	0.349	
High rising (T2)	1	15.083***	9.331***	2.798*	
Low falling (T4)	1	8.227***	5.961***	0.000	
Effects of syllable positions (initial vs. medial vs. final)					
	Kruskal–Wallis test (*H*)				
	df	AOS	nAOS	HC	
High level (T1)	2	4.252	75.549***	2.000	
High rising (T2)	2	48.323***	155.403***	75.154***	
Low falling (T4)	2	196.880***	8.367*	0.000	
	Multiple comparison[a]				
	df	Syllable position	AOS	nAOS	HC
High level (T1)	2	Initial vs. medial	NA	5.237***	NA
	2	Initial vs. final	NA	−8.626***	NA
	2	Medial vs. final	NA	−3.389***	NA
High rising (T2)	2	Initial vs. medial	−3.593**	−3.685**	−3.158**
	2	Initial vs. final	6.950***	12.156***	8.571***
	2	Medial vs. final	3.357**	8.471***	5.413***
Low falling (T4)	2	Initial vs. medial	4.101***	0.823	NA
	2	Initial vs. final	9.570***	−2.813*	NA
	2	Medial vs. final	13.672***	−1.990	NA

Note. V = vowel structure; CV = consonant–vowel structure; NA = not applicable.

[a]Bonferroni correction was applied.

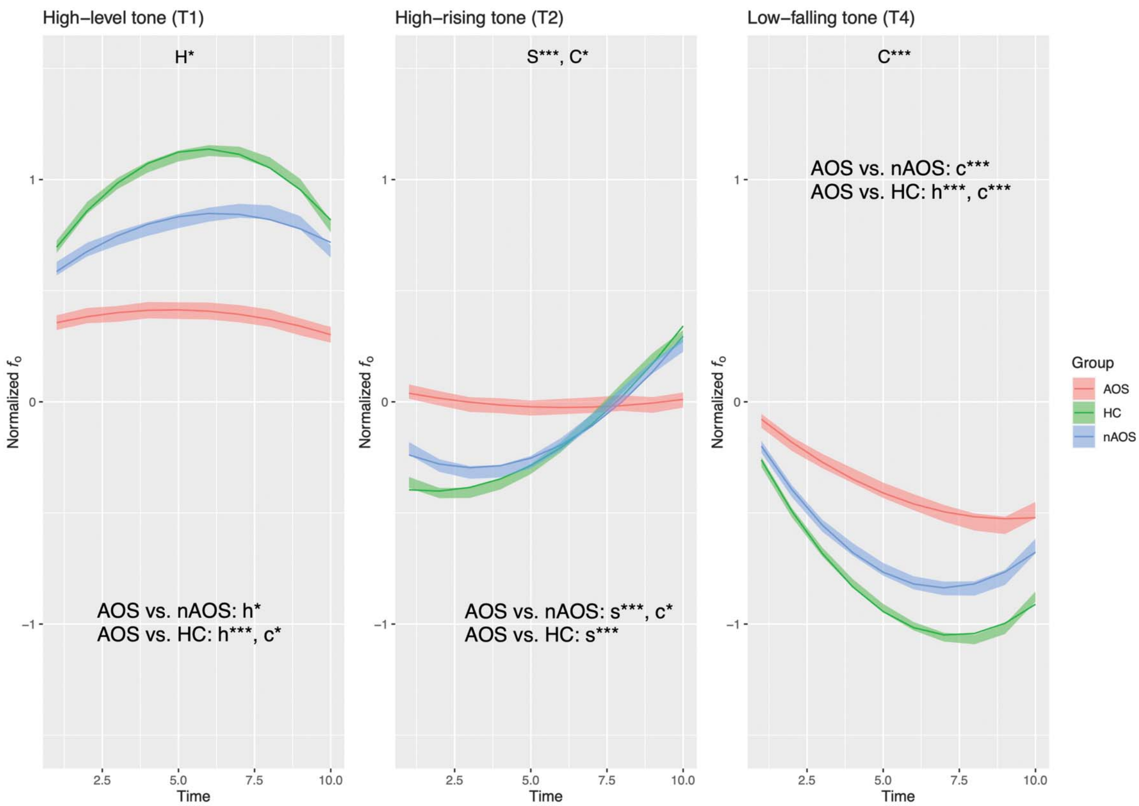
*$p < .05$. **$p < .01$. ***$p < .001$.

Tone accuracy of T1, T2, and T4 was significantly different among the three groups, with only the HC group producing all of the T1 and T4 syllables correctly (i.e., ceiling performance). In particular, the AOS group had the worst performance, followed by the nAOS and HC groups. The participants in the AOS group produced T1 with the highest accuracy, followed by T4 and T2. The descriptive statistics are presented in Appendix E.

f_o Values

The model fits for normalized f_o values by group are shown in Figure 1 and Appendix D. For T1 (high-level) syllables, the groups were significantly different on pitch height only, $\chi^2(2) = 8.986$, $p < .05$. The AOS group showed significantly lower pitch heights than the nAOS and HC groups, while the nAOS and HC groups had comparable pitch heights. For T2 (high-rising) syllables, the groups were significantly different on pitch slope, $\chi^2(2) = 14.771$, $p < .001$, and curvature, $\chi^2(2) = 9.064$, $p < .05$. The AOS group produced less steep pitch slopes compared with the nAOS and HC groups, while the control groups did not differ from each other. It was observed that the AOS group did not produce a rising contour for T2 syllables. For T4 (low-falling) syllables, the groups were

Figure 1. Observed data and growth curve model fits for the effect of group on normalized fundamental frequency values in syllables with high-level (T1; left), high-rising (T2; middle), and low-falling (T4; right) tones. The uppercase letters indicate the significance of model fit: H = pitch height; S = pitch slope; C = pitch curvature. The lowercase letters indicate the significance of parameter estimates: h = pitch height; s = pitch slope; c = pitch curvature. *$p < .05$, **$p < .01$, ***$p < .001$. AOS = apraxia of speech; nAOS = no apraxia of speech; HC = healthy control.



significantly different on pitch curvature only, $\chi^2(2) = 19.781$, $p < .001$. The AOS group produced less curved pitch contours compared with the nAOS and HC groups, while the control groups did not differ from each other. It was observed that the AOS group produced a falling but less curved contour for T4 syllables. The parameter estimates for normalized f_o values by group are shown in Appendix E, Table E1.

Acoustic Duration

The results of the Kruskal–Wallis test and pairwise comparisons are presented in Table 5. The AOS group produced significantly longer durations than the HC groups. The repetition durations were similar between the AOS and nAOS groups on any sets. The means and standard deviations of repetition duration in the three groups of participants are presented in Appendix G.

Effects of Syllable Structure

Perceptual Judgment

Table 4 presents the results of Mann–Whitney *U* tests of the comparisons of perceptual judgments between the V and CV structures on the three Cantonese tones in the groups. The AOS and nAOS groups produced T2 syllables with CV structure at a significantly higher accuracy level than those with V structure. The nAOS group produced T1 syllables with V structure with significantly more accuracy than those with CV structure. The HC group produced all tone syllables with different syllable structures with similar tone accuracy levels. The descriptive statistics are presented in Appendix F.

f_o Values

The model fits for normalized f_o values by syllable structure are shown in Figure 2 and Appendix D. For T1 syllables, all of the groups produced items with V and CV structures with significant pitch height differences. The AOS and HC groups showed significant pitch slope differences; the HC group also showed significant pitch curvature differences between items with V and CV structures. It was observed that the AOS group produced higher pitch heights on items with CV structure, while the nAOS and HC groups produced higher pitch heights on items with V structure. Although both AOS and HC groups produced significant pitch slope differences between items

Table 5. Results of Kruskal–Wallis and multiple comparisons for repetition duration in adult speakers with apraxia of speech poststroke (AOS group), adult speakers without apraxia of speech poststroke (nAOS), and healthy controls (HC group).

Repetition duration (s)	Kruskal–Wallis test (H)	Multiple comparisons[a]		
		AOS vs. HC	AOS vs. nAOS	nAOS vs. HC
First set	24.001**	3.651*	–0.692	–4.647**
Second set	396.569**	17.564**	0.655	–16.909**
Third set	67.696**	7.149**	0.031	–7.093**
Overall	243.965**	13.653**	0.254	–13.395**
Syllable structure (V vs. CV)	Mann–Whitney U test			
	AOS	nAOS		HC
First set	3,439.0*	5,775.5*		5,185.0**
Second set	25,181.5	28,653.0**		30,975.0**
Overall	44,240.5	47,229.5*		50,480.5*
Lexical status (words vs. nonwords)	Mann–Whitney U test			
	AOS	nAOS		HC
First set	2,540.5	5,115.5		4,047.0
Third set	622.0	618.0		535.0
Overall	5,644.5	5,966.5		6,211.5

Note. V = vowel structure; CV = consonant–vowel structure.
[a]Bonferroni correction was applied.
*$p < .01$. **$p < .001$.

with V and CV structures, it was observed that the AOS group did not produce steeper pitch slopes on items with V structure as observed in the HC group. Neither the AOS nor the nAOS group produced more curved pitch on items with V structure as observed in the HC group. For T2 syllables, all of the groups produced items with V and CV structures with significant pitch slope and curvature differences. Pitch height differences on items with V versus CV structures were found in the nAOS and HC groups. It was observed that the AOS group did not produce higher pitch height on items with V structure than those with CV structure, as observed in the nAOS and HC groups. The AOS group also did not produce steeper pitch slopes nor more curved pitch as observed in the control groups. Instead, the AOS group produced items with CV structure with a slightly falling pitch and items with V structure with minimal pitch change. For T4 syllables, the nAOS and HC groups had significant pitch height differences on items with V and CV structures, while significant pitch slope and curvature differences were only found in the HC group. It was observed that the AOS group did not produce higher pitch height on items with V structure, nor steeper pitch slopes and more curved pitch on items with CV structure, as observed in the control groups. The parameter estimates for normalized f_o values by syllable structure are shown in Appendix E, Table E2.

Acoustic Duration

The results of Mann–Whitney U tests for the comparisons between the items with V and CV structures on the first and second sets and the overall TST task in each group are presented in Table 5. For the HC and nAOS groups, significant repetition duration differences between items with V and CV structures were found in both the first and second sets as well as the overall TST task. Specifically, items with CV structure were produced with significantly shorter duration than those with V structure. For the AOS group, significant repetition duration differences between items with V and CV structures were only found in the first set, in which items with CV structure were produced with significantly shorter duration than items with V structure. The means and standard deviations of repetition duration of the items with V and CV structures in the three groups in the first and second sets and the overall TST task are presented in Appendix G.

Effects of Lexical Status

Perceptual Judgment

Table 4 presents the results of Mann–Whitney U tests of the perceptual judgments of tone accuracy with different lexical statuses in the groups. The AOS and nAOS groups produced word items with significantly more accuracy than nonword items on all tone syllables. The HC group also produced word items with significantly higher accuracy levels than nonword items, but only in T2 syllables. The descriptive statistics are presented in Appendix F.

f_o Values

The model fits for normalized f_o values by lexical status are shown in Figure 3 and Appendix D. For T1

Figure 2. Observed data and growth curve model fits for the effect of syllable structure on normalized fundamental frequency values in syllables with high-level (T1; first row), high-rising (T2; second row), and low-falling (T4) tones in the groups of adult speakers with apraxia of speech poststroke (AOS group; left), adult speakers without apraxia of speech poststroke (nAOS group; middle), and healthy controls (HC group; right). The uppercase letters indicate the significance of model fit: H = pitch height; S = pitch slope; C = pitch curvature. *p < .05, **p < .01, ***p< .001. CV = consonant–vowel; V = vowel.

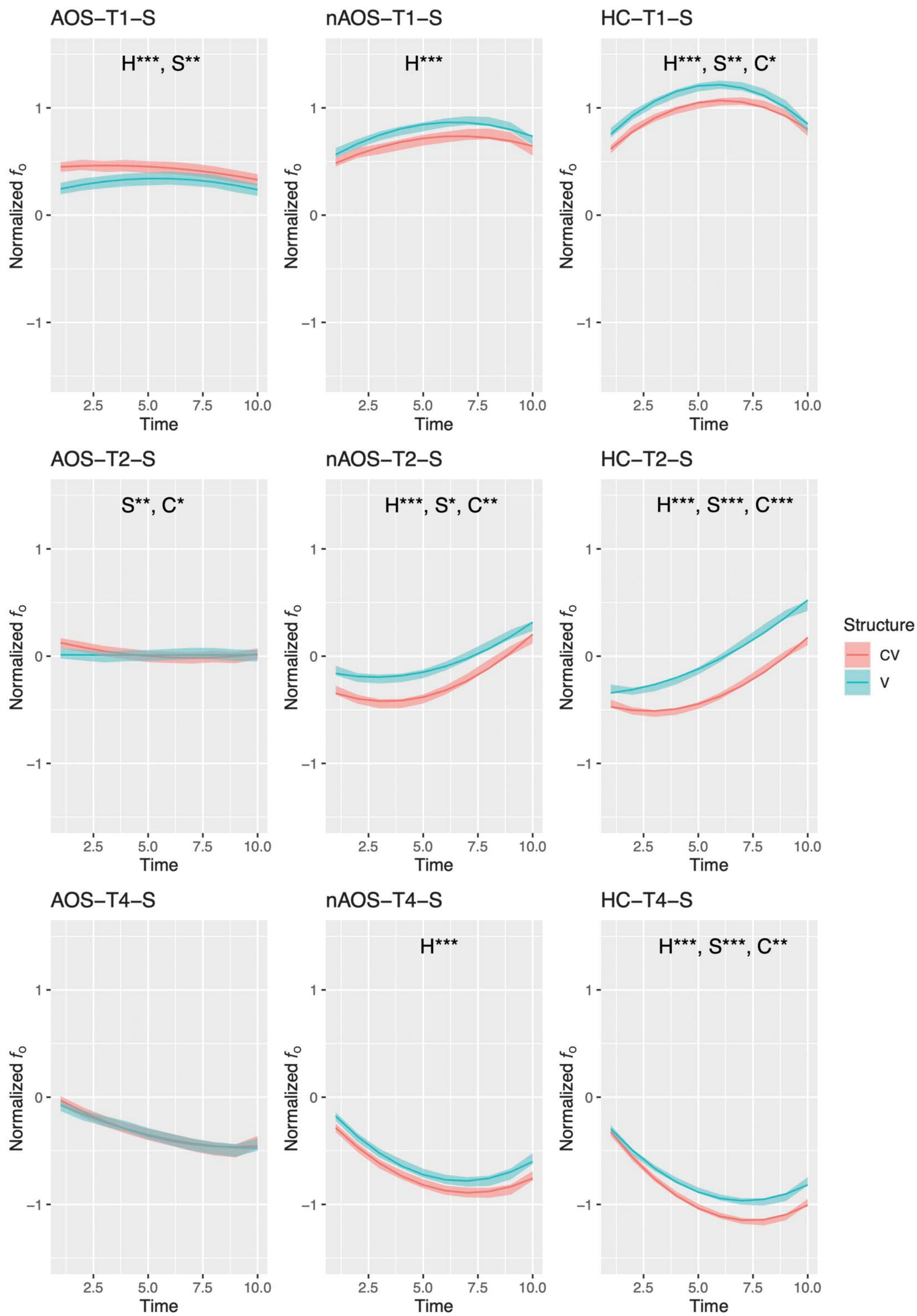
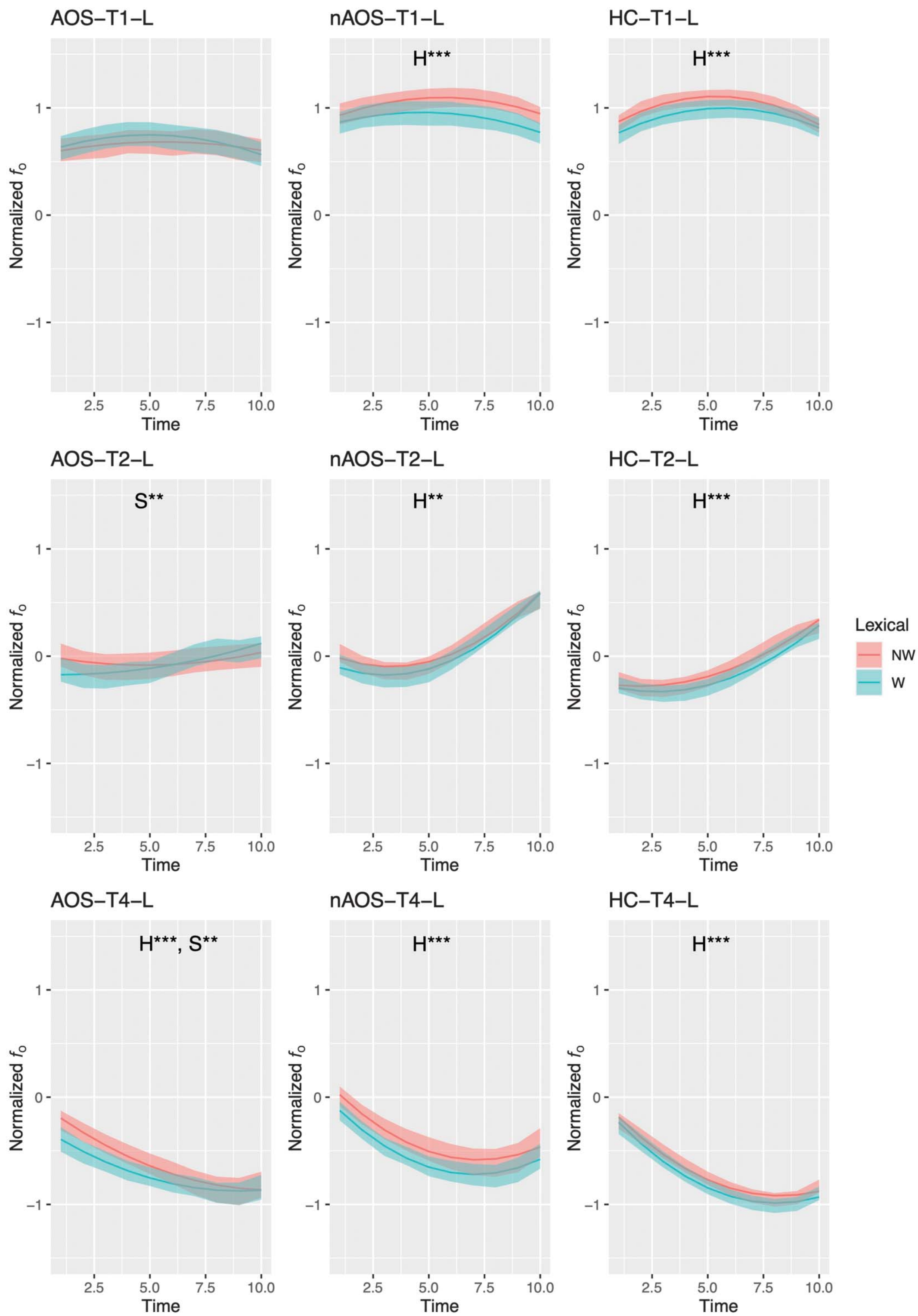


Figure 3. Observed data and growth curve model fits for the effect of lexical status on normalized fundamental frequency values in syllable with high-level (T1; first row), high-rising (T2; second row), and low-falling (T4) tones in the groups of adult speakers with apraxia of speech poststroke (AOS group; left), adult speakers without apraxia of speech poststroke (nAOS group; middle), and healthy controls (HC group; right). The uppercase letters indicate the significance of model fit: H = pitch height; S = pitch slope; C = pitch curvature. *$p < .05$, **$p < .01$, ***$p < .001$. NW = nonword; W = word.



syllables, the nAOS and HC groups produced word and nonword items with significant pitch height differences, while the AOS group produced word and nonword items similarly in terms of pitch height, slope, and curvature. It was observed that the AOS group did not produce higher pitch heights on nonword items, as observed in the control groups. For T2 syllables, the AOS group produced word and nonword items with significant pitch slope differences, while the nAOS and HC groups produced word and nonword items with significant pitch height differences. It was observed that the AOS group did not produce higher pitch heights on nonword items, as observed in the control groups. Instead, the AOS group produced word and nonword items similarly in terms of pitch height and curvature. For T4 syllables, the AOS, nAOS, and HC groups produced word and nonword items with significant pitch height differences, while the AOS group also produced word and nonword items with significant pitch slope differences. It was observed that the AOS group produced higher pitch height on nonword items, similar to the control groups. The parameter estimates for normalized f_o values by lexical status are shown in Appendix E, Table E2.

Acoustic Duration

The results of Mann–Whitney *U* tests for comparisons between word and nonword items on the first and third sets and the overall TST task in each group are presented in Table 5. There were no significant differences in repetition duration between word and nonword items by any of the groups in any of the sets. The means and standard deviations of repetition duration of word and nonword items in the three groups in the first and third sets and the overall TST tasks are presented in Appendix G.

Effects of Syllable Position

Perceptual Judgment

Table 4 presents the results of the Kruskal–Wallis and post hoc analyses of the perceptual judgments of tone accuracy in the three positions. For T1 syllables, only the nAOS group produced items in different positions with significantly different tone accuracy. Specifically, the highest accuracy was found in the initial position, followed by the medial and final positions. For T2 syllables, all of the groups produced items in different positions with significantly different tone accuracy levels, with the highest accuracy in the final position, followed by the medial and initial positions. For T4 syllables, the AOS group produced items in different positions with significantly different tone accuracy levels, with the highest accuracy in the final position, followed by the initial and medial positions. The nAOS group produced items in the initial position with significantly more accuracy than items in the final position. The HC group showed similar tone accuracy across different positions. The descriptive statistics are presented in Appendix F.

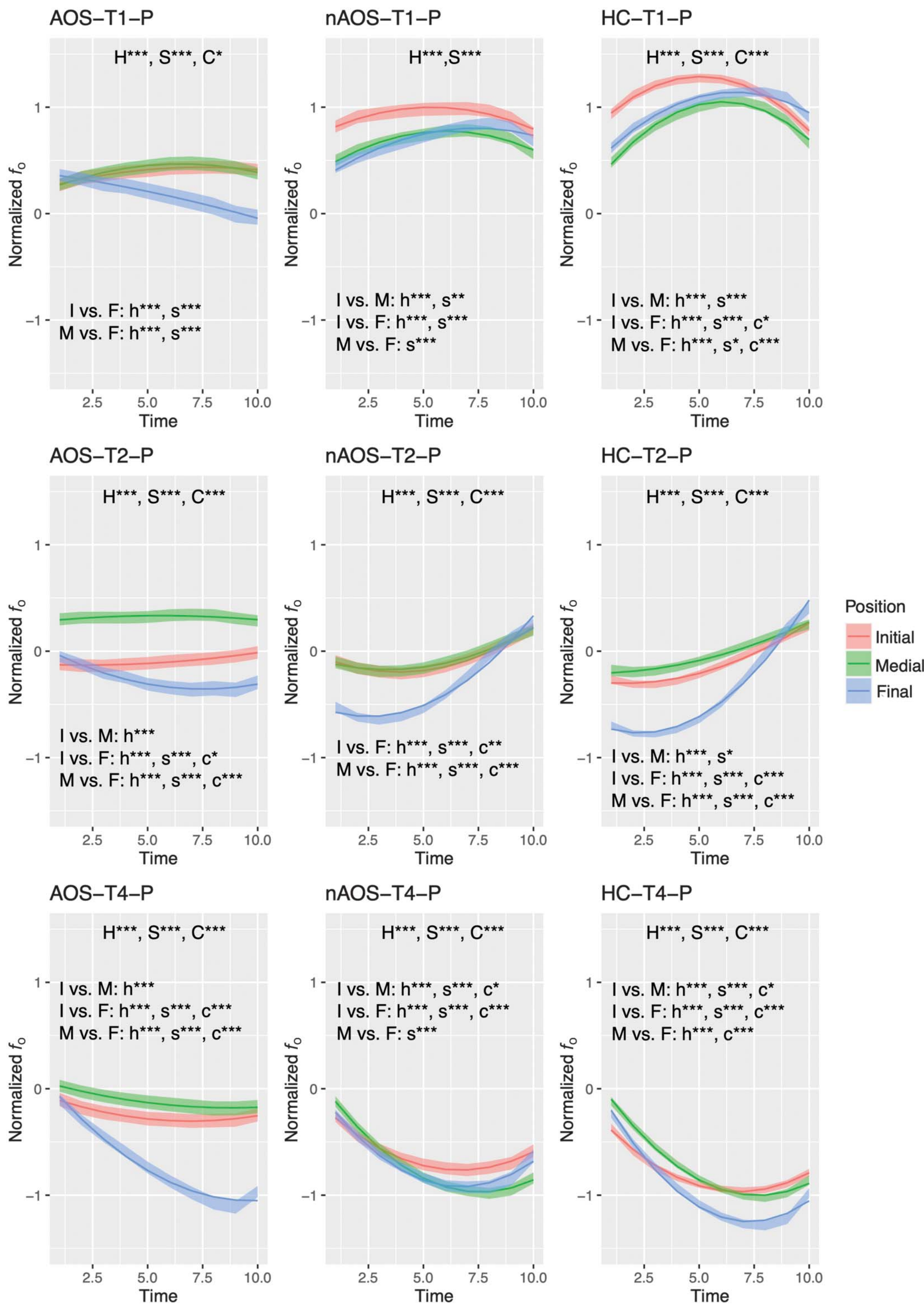
f_o Values

The model fits for normalized f_o values by syllable position are shown in Figure 4 and Appendix D. For T1 syllables, all of the groups produced items in different positions with significant pitch height, slope, and curvature differences, except pitch curvature in the nAOS group. It was observed that the AOS group produced items in the final position with significantly lower pitch height and significantly steeper but falling pitch slope, which were not observed in the control groups. The nAOS and HC groups produced items in the initial position with significantly higher pitch heights, and the HC group produced items in the final position with significantly more curved pitch. For T2 syllables, all of the groups produced items in different positions with significant pitch height, slope, and curvature differences. The AOS group produced items in the medial position with significantly higher pitch heights, which was same as the HC group. The AOS group also produced items in the final position with significantly lower pitch height, steeper pitch slope, and more curved pitch, the same as with the nAOS and HC groups. However, the AOS group produced falling pitch for T2 syllables in the final position, while rising pitch was observed in the control groups. For T4 syllables, all of the groups produced items in different positions with significant pitch height, slope, and curvature differences. The AOS group produced items in the medial position with significantly higher pitch heights, which was the same as the HC group but different from the nAOS group. The AOS group also produced items in the final position with significantly lower pitch height, steeper pitch slope, and more curved pitch, again the same as the HC group but different from the nAOS group. However, the AOS group did not produce items in the initial and medial positions with significantly steeper pitch slope or more curved pitch, which was observed in the nAOS and HC groups. The parameter estimates for normalized f_o values by syllable position are shown in Appendix E, Table E3.

Mixed-Effects Binomial Logistic Regression

The results of model fitting showed that the best models for all three groups and all participants were the same, Accuracy ~ (1| Participant) + Syllable + Position. The model's total explanatory power was substantial in the all participants model (conditional R^2 = .97, marginal R^2 = .92), in the nAOS group model (conditional R^2 = .56, marginal R^2 = .13), and in the HC group model (conditional R^2 = .97, marginal R^2 = .92). The model's total explanatory power was moderate in the AOS group model (conditional R^2 = .21, marginal R^2 = .11).

Figure 4. Observed data and growth curve model fits for the effect of syllable position on normalized fundamental frequency values in syllables with high-level (T1; first row), high-rising (T2; second row), and low-falling (T4) tones in the groups of adult speakers with apraxia of speech poststroke (AOS group; left), adult speakers without apraxia of speech poststroke (nAOS group; middle), and healthy controls (HC group; right). The uppercase letters indicate the significance of model fit: H = pitch height; S = pitch slope; C = pitch curvature. The lowercase letters indicate the significance of parameter estimates: h = pitch height; s = pitch slope; c = pitch curvature. *p < .05, **p < .01, ***p < .001.



The AOS group had significantly higher tone accuracy on items with CV than V structures. The effects of syllable position were statistically significant only in the AOS and nAOS groups. The AOS group had significantly higher tone accuracy on items in the final than initial and medial positions, while the nAOS group had significantly higher tone accuracy on T2 syllables in the final position and on T1 syllables in the initial position. The beta and standardized beta coefficients of the effects of Syllable and Position within each of the best-fitted models are presented in Table 6.

Discussion

Principal Findings

The purpose of this study was to examine the pitch variation skills of adults with AOS poststroke in sequencing context. The first hypothesis, which stated that the f_o values of tone syllables produced by the AOS group would differ from those produced by the control groups, is supported. The results of acoustic analyses showed that the AOS group produced level tone with lower pitch heights than the control groups. For contour tones, the AOS group produced pitch slopes that were less steep and pitch curves that were flatter than those produced by the control groups. These results were inconsistent with that of Tee et al. (2022), which found a substitution of T4 (low-falling) syllables with T2 (high-rising) syllables. Moreover, this study found that speakers with AOS poststroke had more difficulty in producing rising tone than falling tone. The observed difficulty in producing the rising contour of T2 syllables is consistent with the findings reported in children with AOS (E. C. H. Wong et al., 2021) and in Mandarin speakers with AOS (Chen et al., 2022). The low tone accuracy and long repetition duration observed in the AOS group are also consistent with those reported in children with AOS (E. C. H. Wong et al., 2021). The speakers with AOS poststroke were unable to increase their tone accuracy even when they spent more time producing the stimuli. This difficulty could be explained by their deficits in motor planning of speech movements, resulting in difficulty planning of laryngeal muscle movements to achieve different tone contours.

The results of acoustic analyses also showed that the nAOS group had similar tone productions as the HC group in terms of pitch height, slope, and curvature. However, the perceptual results showed that the nAOS group had significantly lower tone accuracy than the HC groups. These findings suggest that acoustic analysis of tone productions in the sequencing context may have the potential to differentiate Cantonese speakers with and without AOS poststroke, but not between Cantonese speakers without AOS and healthy speakers. Additionally, perceptual judgment of tone accuracy in the sequencing context may differentiate Cantonese speakers with aphasia without AOS and healthy speakers.

Effects of Syllable Structure

The second hypothesis, that the AOS group would produce different f_o values on items with different syllable structures while the control groups would not, is partially supported. There were three predictions. The first prediction was related to the f_o values. It posited that the control groups would produce similar f_o values on items with

Table 6. Fitted data of the mixed-effect binomial logistic regression models for the prediction of perceptual judgment of tone accuracy with syllable structure and syllable position effects.

Participant	Predictor	Beta	95% Confidence interval	*p*	Standardized beta	95% Confidence interval
All participants	Syllable structure (V)	−0.35	[−0.73, 0.03]	.070	−0.35	[−0.73, 0.03]
	Syllable position (initial)	−19.95	[−4535.10, 4495.21]	.993	−19.95	[−4535.10, 4495.21]
	Syllable position (medial)	−19.43	[4534.58, 4495.73]	.993	−19.43	[−4534.58, 4495.73]
AOS group	Syllable structure (V)	−0.13	[−0.23, −0.03]	.011*	−0.13	[−0.23, −0.03]
	Syllable position (initial)	−0.64	[0.75, −0.52]	< .001***	−0.64	[−0.75, −0.52]
	Syllable position (medial)	−0.64	[−0.76, −0.52]	< .001***	−0.64	[−0.76, −0.52]
nAOS group	Syllable structure (V)	−0.11	[0.24, 0.03]	.120	−0.11	[−0.24, 0.03]
	Syllable position (initial)	−0.24	[0.41, −0.08]	.004**	−0.24	[−0.41, −0.08]
	Syllable position (medial)	−0.29	[−0.45, −0.12]	< .001***	−0.29	[−0.45, −0.12]
HC group	Syllable structure (V)	−0.35	[0.73, 0.03]	.070	−0.35	[−0.73, 0.03]
	Syllable position (initial)	−19.95	[−4535.10, 4495.21]	.993	−19.95	[−4535.10, 4495.21]
	Syllable position (medial)	−19.43	[−4534.58, 4495.73]	−.993	−19.43	[−4534.58, 4495.73]

Note. The variables of syllable structure are V (vowel) and CV (consonant–vowel) structures; syllable positions are initial, medial, and final. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls.
**p* < .05. ***p* < .01. ****p* < .001.

V and CV structures, whereas the AOS group would exhibit a flatter pitch change on items with CV structure compared with those with V structure. This prediction is not supported by the results because both the nAOS and HC groups produced items with different syllable structures differently. This finding contrasted with those of Xu's (1998) study, which reported similar f_o values across different syllable structures in healthy Mandarin speakers. The discrepancy may be due to methodological differences, as in this study, the speakers were required to produce syllables in a sequencing context. Separating V syllables in a sequence requires higher laryngeal effort (E. C. H. Wong et al., 2021), which may lead to changes in f_o values. The AOS group did not exhibit any pitch height differences on the contour tones, possibly due to their difficulty in producing rising or falling contours, as discussed earlier. However, they showed a different pattern from the control groups on the high-level tone (T1), which was the least challenging tone in this study. Specifically, they produced higher pitch heights on items with CV structure, suggesting that they exerted more effort to produce CV syllables in the sequencing context. Although the prediction is not supported, the results imply that syllable structure may impact pitch variation skills among Cantonese speakers with AOS poststroke.

The second prediction, that the AOS group would have lower tone accuracy on items with CV structure, is not supported. The AOS group produced T1 and T4 syllables with CV and V structures similarly but had a significantly different level of tone accuracy on T2 syllables. The results of logistic regression also confirmed that the effects of syllable structure are present in Cantonese speakers with AOS poststroke. Given that the AOS group had relatively better performance on T1 and T4 syllables compared with T2 syllables, the findings suggest that Cantonese speakers with AOS poststroke are more capable in managing both syllable structure and tone when the tones involved are the relatively easier level or falling tones compared with the more challenging rising tone.

The last prediction, that all of the groups would take longer to repeat items with V structure than CV structure, is partially supported. The lack of differences in the overall repetition duration between items with V and CV structures in the AOS group is inconsistent with the study of E. C. H. Wong et al. (2021), in which children with AOS produced longer repetition duration on items with V than CV structures, the same as the control groups. The dissimilarities may be attributed to differences in oral–laryngeal muscle control between adults and children. It appears that adults possess more refined muscle control than children, as evidenced by the results indicating that adults with AOS poststroke can manage relatively complex syllable structures in some tone syllables, as previously discussed. Further research is necessary to gain a better understanding of this issue.

Effects of Lexical Status

The third hypothesis that the effects of lexical status would be found in both the AOS and control groups is supported. The similar repetition duration found between word and nonword items is unsurprising based on the findings from Cantonese-speaking children with AOS (E. C. H. Wong et al., 2021). The results of acoustic and perceptual analyses showed that the lexical status facilitates tone production in the AOS group as a priming effect. The significantly steeper pitch slope on word items with T4 (low-falling) and T2 (high-rising) syllables induced a significantly higher perceptual tone accuracy in the AOS group, compared with the nonword items. The results are inconsistent with the findings from Cantonese-speaking children with AOS, in which no priming effect was found (E. C. H. Wong et al., 2021). The different impacts of priming effects in Cantonese-speaking children and adults with AOS might be explained by differences between developing and developed semantic inventories. Finally, the findings suggest that the effects of word versus nonword items in our study are somewhat analogous to the high- versus low-frequency words observed by Zhao and Jurafsky (2009). Specifically, the significantly higher pitch heights observed on nonword items in our participants are consistent with the results of Zhao and Jurafsky's investigation, in which Cantonese speakers produced low-frequency words with higher pitch heights as compared to high-frequency words.

Effects of Tone Syllable Position

The final hypothesis, that there would be more f_o differences among the syllables in three different syllable positions in the AOS group and fewer f_o differences in the control groups, is not supported. The AOS group exhibited a comparable number of differences as the nAOS group and fewer differences than the HC group. These results expand the understanding of tonal co-articulation effects in Cantonese speakers. Even when the tones of neighboring syllables are controlled as in the context of the TST, pitch performance (in terms of pitch height, slope, and curvature) still varied among the syllable positions, especially in healthy adults' speech. The findings are consistent with those of Fon and Hsu (2004), in which tone contours also varied by position.

For speakers with AOS poststroke, fewer differences were observed across the three different tones in three tone positions. This further suggests that they have degraded pitch variation skills compared with the controls. Another finding was that the f_o change patterns produced by the AOS group differed from those of the control groups, while the two control groups demonstrated similar patterns. The AOS group produced a unique lowering of f_o values in the final position, which was not observed in any of the control groups. Perceptually, the patterns were detected as incorrect for high-level (T1) and high-rising (T2) tones but were deemed correct for the low-falling tone (T4).

In addition, although actual acoustic differences were found among different syllable positions, the results of perceptual judgments showed that some f_o differences were not detected by the listeners. In the AOS group, the pitch height and slope differences on T1 syllables between the initial versus final and medial versus final position in the AOS group were not detected. In the HC group, the pitch height, slope, and curvature differences on T1 and T4 syllables across the three syllable positions were not detected. In the nAOS group, the pitch height, slope, and curvature differences on T4 syllables between initial versus medial positions and the pitch slope differences on T4 syllables between medial versus final positions were also not detected by the listeners.

In summary, the results of this study suggest that adults with AOS poststroke may have degraded pitch variation skills compared with those without AOS poststroke and healthy controls. Linguistic variables such as complexity of syllable structure, lexical status, and syllable position affect pitch variation skills in adults with AOS poststroke.

Clinical Implication

The results of this study suggest that the TST could differentiate adults with AOS poststroke from those without AOS poststroke and healthy controls. Acoustic pitch characteristics, including pitch height, slope, and curvature, as well as the perceptual accuracy of T1 (high-level), T2 (high-rising), and T4 (low-falling) syllables in a sequential context, may differentiate between Cantonese speakers with and without AOS poststroke. These differences are particularly evident in the performance of T2 syllables with varying syllable structures and lexical status. Furthermore, individuals with AOS poststroke may exhibit a reduction in f_o values in the final position when repeating three-syllable items. Although this study showed that repetition durational measures could differentiate speakers with and without AOS poststroke, application of this measure in clinical settings requires extra caution, given that some participants in the AOS and nAOS groups were diagnosed with co-existing mild dysarthria (e.g., flaccid and hypokinetic dysarthria). Syllable repetition rate in hypokinetic dysarthria could be faster, slower, or the same as in healthy speakers (M. N. Wong et al., 2012), although Kent et al. (2022) suggested that slower rate is more common. Regarding other types of dysarthria, Hartman and Abbs (1992) reported that individuals with unilateral upper motor neuron lesions had slow and irregular DDK, and Thompson et al. (1995) observed slower rates of repetitive lip movements in individuals with upper motor neuron dysarthria. For speakers with flaccid dysarthria, Duffy (2020) suggested that they have slower or normal syllable repetition rates. However, it is unknown if these types of dysarthria play a role in tone syllable repetition tasks as there has been no such investigation to date. The presence of dysarthria, although mild, may have confounded the results. The results also suggest that the TST can be further developed as a diagnostic tool for adults with AOS poststroke. However, the large number of items is clinically unfriendly. Further investigations are recommended to determine the effectiveness of individual items in differentiating adults with and without AOS poststroke.

Limitations and Future Investigations

This study has several limitations, and other issues remain to be investigated further. First, regarding syllable structure, only CV and V structures were analyzed. Items with the most complex syllable structure in Cantonese (i.e., CVC) were neither controlled nor analyzed in this study. Future investigations involving other relatively complex syllable structures are needed to further examine the effects of syllable structure on pitch variation skills. Second, although the three Cantonese tones (i.e., T1, T2, and T4) represent the high and low tones and three different contours in Cantonese tones, the midlevel tone (i.e., T3) has not been included in the TST. Future investigations of pitch variation may include analyses of this midlevel tone. Another limitation is that some participants in the AOS group showed impaired tone identification skills, and this may have affected their tone production in this study. Some differences between the AOS and control groups might be attributable to their differences in tone perception ability. Future investigations should control the participants' tone perception ability levels. Finally, as mentioned, the presence of co-existing mild dysarthria in some of the participants in the AOS and nAOS groups may have confounded the findings. Future investigations with well-controlled participants are recommended. Future studies may also adopt the TST in Mandarin and other tonal languages to examine pitch variation skills in other tonal language speakers with AOS poststroke.

Conclusions

This study was the first ever investigation that examined the pitch variation skills and potentially contributing linguistic factors in Cantonese-speaking adults with AOS poststroke. The results suggest that adults with AOS poststroke may show degraded pitch variation skills compared with those without AOS poststroke and healthy controls. Syllable structure, lexical status, and syllable position appear to affect pitch variations in adults with AOS poststroke. The TST is suggested to be a potential diagnostic tool to differentiate adults with AOS poststroke from those without AOS and healthy controls. Future investigations may adopt the TST in Mandarin and other tonal languages to investigate pitch variation skills in other tonal language speakers with AOS.

Data Availability Statement

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

Acknowledgments

This study was funded by the Dean's Reserve for Research, Scholarly and Other Endeavours (Project ID P0036018) of the Faculty of Humanities, the Hong Kong Polytechnic University, to the first and second authors. The authors expressed their gratitude to all of the participants and their families for their participation. The authors would also like to thank the student helpers and research assistants who performed the data analyses and reliability measures, including Twinkie Cheung, Helen Chung, Jenna Mak, Carly Ng, Ventus Siu, Dahlia Tang, Sicilia Wong, and Cherry Yu. Appreciation is also due to the speech-language pathology student, Zenas Chan, for his participation in data collection and to two SLPs, Wilbert Chan and Jason Kan, for their viewing of the assessment videos and making diagnoses.

References

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

Bakhtiar, M., Wong, M. N., Tsui, E. K. Y., & McNeil, M. R. (2020). Development of the English listening and reading computerized revised token test into Cantonese: Validity, reliability, and sensitivity/specificity in people with aphasia and healthy controls. *Journal of Speech, Language, and Hearing Research, 63*(11), 3743–3759. https://doi.org/10.1044/2020_JSLHR-20-00103

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G., & van Doorn, J. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*(6), 1822–1835. https://doi.org/10.1044/1092-4388(2012/11-0257)

Ballard, K. J., Savage, S., Leyton, C. E., Vogel, A. P., Hornberger, M., & Hodges, J. R. (2014). Logopenic and nonfluent variants of primary progressive aphasia are differentiated by acoustic measures of speech production. *PLOS ONE, 9*(2), Article e89864. https://doi.org/10.1371/journal.pone.0089864

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

Chen, W., van de Weijer, J., Qian, Q., Zhu, S., & Wang, M. (2022). Tone and vowel disruptions in Mandarin aphasia and apraxia of speech. *Clinical Linguistics & Phonetics, 37*(8), 742–765. https://doi.org/10.1080/02699206.2022.2081611

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Duffy, J. R., Utianski, R. L., & Josephs, K. A. (2020). Primary progressive apraxia of speech: From recognition to diagnosis and care. *Aphasiology, 35*(4), 560–591. https://doi.org/10.1080/02687038.2020.1787732

Enderby, P., & Palmer, R. (2008). *Frenchay Dysarthria Assessment–Second Edition*. Pro-Ed.

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

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

Haley, K. L., & Jacks, A. (2019). Word-level prosodic measures and the differential diagnosis of apraxia of speech. *Clinical Linguistics & Phonetics, 33*(5), 479–495. https://doi.org/10.1080/02699206.2018.1550813

Hartman, D. E., & Abbs, J. H. (1992). Dysarthria associated with focal unilateral upper motor neuron lesion. *International Journal of Language and Communication Disorders, 27*(3), 187–196. https://doi.org/10.3109/13682829209029419

Kent, R. D., Kim, Y., & Chen, L. M. (2022). Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications. *Journal of Speech, Language, and Hearing Research, 65*(2), 574–623. https://doi.org/10.1044/2021_jslhr-21-00396

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.

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

Lee, K. Y. S., Chiu, S. N., & van Hasselt, C. A. (2002). Tone perception ability of Cantonese-speaking children. *Language and Speech, 45*(Pt. 4), 387–406. https://doi.org/10.1177/00238309020450040401

Lennes, M. (2002). *Calculate_segment_durations* [Praat script].

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
Luke, K. K., & Wong, M. L. Y. (2015). The Hong Kong Cantonese corpus: Design and uses. *Journal of Chinese Linguistics Monograph Series, 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
McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 249–268). Thieme.
Miller, H. E., & Guenther, F. H. (2020). Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology, 35*(4), 424–441. https://doi.org/10.1080/02687038.2020.1765307
Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.
Pierce, J. E., Cotton, S., & Perry, A. (2013). Alternating and sequential motion rates in older adults. *International Journal of Language & Communication Disorders, 48*(3), 257–264. https://doi.org/10.1111/1460-6984.12001
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
Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The Apraxia of Speech Rating Scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders, 51,* 43–50. https://doi.org/10.1016/j.jcomdis.2014.06.008
Tee, B. L., Deleon, J., Chen Li Ying, L. K., Miller, B. L., Lo, R. Y., Europa, E., Sudarsan, S., Grasso, S., & Gorno-Tempini, M. L. (2022). Tonal and orthographic analysis in a Cantonese-speaking individual with nonfluent/agrammatic variant primary progressive aphasia. *Neurocase, 28*(1), 1–10. https://doi.org/10.1080/13554794.2021.1925302
Terband, H., Namasivayam, A., Maas, E., van Brenk, F., Mailend, M. L., Diepeveen, S., van Lieshout, P., & Maassen, B. (2019). Assessment of childhood apraxia of speech: A review/tutorial of objective measurement techniques. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2999–3032. https://doi.org/10.1044/2019_jslhr-s-csmc7-19-0214
Thompson, E. C., Murdoch, B. E., & Stokes, P. D. (1995). Lip function in subjects with upper motor neuron type dysarthria following cerebrovascular accidents. *European Journal of Disorders of Communication, 30*(4), 451–466. https://doi.org/10.3109/13682829509087244
Utianski, R. L., Duffy, J. R., Clark, H. M., Strand, E. A., Botha, H., Schwarz, C. G., Machulda, M. M., Senjem, M. L., Spychalla, A. J., Jack, C. R. Jr., Petersen, R. C., Lowe, V. J., Whitwell, J. L., & Josephs, K. A. (2018). Prosodic and phonetic subtypes of primary progressive apraxia of speech. *Brain and Language, 184,* 54–65. https://doi.org/10.1016/j.bandl.2018.06.004
Vergis, M. K., Ballard, K. J., Duffy, J. R., McNeil, M. R., Scholl, D., & Layfield, C. (2014). An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology, 28*(5), 554–575. https://doi.org/10.1080/02687038.2014.889275
Wambaugh, J. L., Duffy, J. R., McNeil, M., Robin, D., & Rogers, M. A. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology, 14*(2), 15–34.
Whitehill, T. L., Ma, J. K., & Lee, A. S. (2003). Perceptual characteristics of Cantonese hypokinetic dysarthria. *Clinical Linguistics & Phonetics, 17*(4–5), 265–271. https://doi.org/10.1080/0269920031000080082
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., Velleman, S. L., Tong, M. C. F., & Lee, K. Y. S. (2022). 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
Wong, M. N., Murdoch, B. E., & Whelan, B. M. (2012). Lingual kinematics during rapid syllable repetition in Parkinson's disease. *International Journal of Language and Communication Disorders, 47*(5), 578–588. https://doi.org/10.1111/j.1460-6984.2012.00167.x
Xu, Y. (1998). Consistency of tone syllable alignment across different syllable structures and speaking rates. *Phonetica, 55*(4), 179–203. https://doi.org/10.1159/000028432
Xu, Y. (2013). ProsodyPro—A tool for large-scale systematic prosody analysis. *Tools and resources for the analysis of speech prosody (TRASP 2013)*.
Yip, V., & Matthews, S. (2011). *Cantonese: A comprehensive grammar* (2nd ed.). Routledge.
Yiu, E. M. L. (1992). Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery. *Journal of Neurolinguistics, 7*(4), 379–424. https://doi.org/10.1016/0911-6044(92)90025-R
Zhang, J. (2018). *A comparison of tone normalization methods for language variation research*. Proceedings of the 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

Presence of Characteristics Determined by Two Independent Speech-Language Pathologists Using Apraxia of Speech Rating Scale 3.0

Items	Features	Participant ID					
		1	2	5	10	12	17
Phonetic features							
1	Sound distortions	✓	✓	✓	✓	✓	✓
2	Distorted sound substitutions	✓	✓	✓	✓	✓	✓
3	Distorted sound additions	✓	✓	✓	✓	✓	
4	Increased errors with increased utterance length or syllable/word articulatory complexity	✓		✓	✓	✓	✓
Prosodic features							
5	Syllable segmentation within words	✓	✓	✓	✓	✓	✓
6	Syllable segmentation across words	✓	✓	✓	✓	✓	✓
7	Slow overall speech rate	✓		✓	✓	✓	✓
8	Lengthened vowel and/or consonant segments	✓	✓	✓	✓		
Other							
9	Off-target in AMRs	✓	✓	✓	✓	✓	✓
10	Deliberate, slowly sequenced, segmented, and/or distorted in SMRs	✓	✓	✓	✓	✓	✓
11	Consistently reduced words per breath group						
12	Visible or silent articulatory groping	✓	✓	✓			✓
13	Audible false starts/restarts	✓	✓	✓	✓		✓

Note. AMRs = alternate motion rates; SMRs = sequential motion rates. ASRS items reprinted from Utianski et al. (2018), Copyright © 2018, with permission from Elsevier.

Appendix B (p. 1 of 3)

Items of the Tone Sequencing Task

Table B1. Items included in the first set of the tone-sequencing task.

Variable	Word stimuli	Meaning	Transcription	Nonword stimuli with transcription
V structure				
Tone 1	呀	Particle for exclamation	[a1]	[ɛ1]
Tone 2	啞	Dumb	[a2]	[ɛ2]
Tone 4	牙	Tooth/teeth	[a4]	[ɛ4]
Tone 1	屙	Excretion	[ɔ1]	[i1]
Tone 2	鵝	Goose	[ɔ2]	[i2]
Tone 4	鵝	Goose	[ɔ4]	[i3]
CV structure				
Tone 1	咩	Sound of sheep or question word	[mɛ1]	[jɛ1]
Tone 2	歪	Crooked	[mɛ2]	[jɛ2]
Tone 4	爺	Grandfather	[jɛ4]	[mɛ4]
Tone 1	多	More	[tɔ1]	[nɔ1]
Tone 2	朵	Classifier for flower	[tɔ2]	[nɔ2]
Tone 4	挪	Take	[nɔ4]	[tɔ4]

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

Appendix B (p. 2 of 3)

Items of the Tone Sequencing Task

Table B2. Items included in the second set of the tone-sequencing task.

		Second syllable									
		[a1]/ [mɛ1]	[a2]/ [mɛ2]	[a4]/ [mɛ4]	[ɔ1]/ [jɛ1]	[ɔ2]/ [jɛ2]	[ɔ4]/ [jɛ4]	[ɛ1]/ [tɔ1]	[ɛ2]/ [tɔ2]	[ɛ4]/ [tɔ4]	Third syllable
First syllable	[a1]/ [mɛ1]	■	■	■	■	■	■	■	■		[ɔ2]/ [jɛ2]
		■	■	■	■	■	■	■		■	[ɔ4]/ [jɛ4]
		■	■	■	■	■		■	■	■	[ɛ2]/ [tɔ2]
		■	■	■	■		■	■	■	■	[ɛ4]/ [tɔ4]
	[a2]/ [mɛ2]	■	■	■	■	■	■	■	■		[ɔ1]/ [jɛ1]
		■	■	■	■	■	■		■	■	[ɔ4]/ [jɛ4]
		■	■	■	■	■		■	■	■	[ɛ1]/ [tɔ1]
		■	■	■		■	■	■	■	■	[ɛ4]/ [tɔ4]
	[a4]/ [mɛ4]	■	■	■	■	■	■	■		■	[ɔ1]/ [jɛ1]
		■	■	■	■	■	■		■	■	[ɔ2]/ [jɛ2]
		■	■	■	■		■	■	■	■	[ɛ1]/ [tɔ1]
		■	■	■		■	■	■	■	■	[ɛ2]/ [tɔ2]
	[ɔ1]/ [jɛ1]	■	■	■	■	■	■	■	■		[a2]/ [mɛ2]
		■	■	■	■	■	■	■		■	[a4]/ [mɛ4]
		■	■		■	■	■	■	■	■	[ɛ2]/ [tɔ2]
		■		■	■	■	■	■	■	■	[ɛ4]/ [tɔ4]
	[ɔ2]/ [jɛ2]	■	■	■	■	■	■	■	■		[a1]/ [mɛ1]
		■	■	■	■	■	■		■	■	[a4]/ [mɛ4]
		■	■		■	■	■	■	■	■	[ɛ1]/ [tɔ1]
			■	■	■	■	■	■	■	■	[ɛ4]/ [tɔ4]
	[ɔ4]/ [jɛ4]	■	■	■	■	■	■	■		■	[a1]/ [mɛ1]
		■	■	■	■	■	■		■	■	[a2]/ [mɛ2]
		■		■	■	■	■	■	■	■	[ɛ1]/ [tɔ1]
			■	■	■	■	■	■	■	■	[ɛ2]/ [tɔ2]
	[ɛ1]/ [tɔ1]	■	■	■	■	■		■	■	■	[a2]/ [mɛ2]
		■	■	■	■		■	■	■	■	[a4]/ [mɛ4]
		■	■		■	■	■	■	■	■	[ɔ2]/ [jɛ2]
		■		■	■	■	■	■	■	■	[ɔ4]/ [jɛ4]
	[ɛ2]/ [tɔ2]	■	■	■	■	■		■	■	■	[a1]/ [mɛ1]
		■	■	■		■	■	■	■	■	[a4]/ [mɛ4]
		■	■		■	■	■	■	■	■	[ɔ1]/ [jɛ1]
			■	■	■	■	■	■	■	■	[ɔ4]/ [jɛ4]
	[ɛ4]/ [tɔ4]	■	■	■	■		■	■	■	■	[a1]
		■	■	■		■	■	■	■	■	[a2]/ [mɛ2]
		■		■	■	■	■	■	■	■	[ɔ1]/ [jɛ1]
			■	■	■	■	■	■	■	■	[ɔ2]/ [jɛ2]

Note. A colored box indicates that no stimuli can be formed by the vowel (V) or consonant–vowel (CV) structure. A blank box indicates that there is one stimulus formed using the V structure and one stimulus formed using the CV structure.

Table B3. Items included in the third set of the tone-sequencing task.

Nonword stimuli	Transcription	Word stimuli	Meaning	Transcription
牙邊房	[a4 pin1 fɔŋ2]	牙膏座	Toothpaste holder	[a4 kou1 tsɔ2]
工子爐	[kʊŋ1 tsi2 lou4]	哈比人	Hobbit	[ha1 pei2 jɐn4]
打人汁	[ta2 jɐn4 tsɐp1]	搶銀包	To snatch wallet	[tsʰœŋ2 ɐn4 pau1]
提比包	[tʰɐi4 pei2 pau1]	提子汁	Grape juice	[tʰɐi4 tsi2 tsɐp1]
哈銀座	[ha1 ɐn4 tsɔ2]	工人房	Worker's room	[kʊŋ1 jɐn4 fɔŋ2]
搶膏人	[tsʰœŋ2 kou1 jɐn4]	打邊爐	Eat hot pot	[ta2 pin1 lou4]

Appendix C (p. 1 of 2)

R Code Used in the Analyses

Handling missing data

```
> Library(VIM)
> # checking missing data
> mice_plot <- aggr(all, col=c('navyblue','yellow'), numbers=TRUE, sortVars=TRUE, labels=names(all), cex.axis=.7, gap=3,
ylab=c("Missing data","Pattern"))
> Library(mice)
> # imputation by using predictive mean matching
> imputed_Dataall <- mice(all, method = 'pmm')
> completeDataall <- complete(imputed_Dataall)
```

Analysis of *f0* values of T1 syllables between the groups

```
> Library(lme4)
> # create second-order orthogonal polynomial
> t1 <- poly(unique(completeDataT1$Time), 2)
> # create orthogonal polynomial time variables in data frame
> completeDataT1[,paste("ot", 1:2, sep = "")] <- t1[completeDataT1$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
```

Multiple comparisons of T1 syllable in different positions in the AOS group

```
> contrast.matrix = rbind("1st vs. 2nd" = c(0, 0, 0, 1, 0, 0, 0, 0, 0) , "1st vs. 3rd" = c(0, 0, 0, 0, 1, 0, 0, 0, 0) , "2nd vs. 3rd" = c(0, 0, 0,
–1, 1, 0, 0, 0, 0) , "Slope: 1st vs. 2nd" = c(0, 0, 0, 0, 0, 1, 0, 0, 0) , "Slope: 1st vs. 3rd" = c(0, 0, 0, 0, 0, 0, 1, 0, 0) , "Slope: 2nd vs.
3rd" = c(0, 0, 0, 0, 0, –1, 1, 0, 0) , "Degree: 1st vs. 2nd" = c(0, 0, 0, 0, 0, 0, 0, 1, 0) , "Degree: 1st vs. 3rd" = c(0, 0, 0, 0, 0, 0, 0, 0, 1) ,
"Degree: 2nd vs. 3rd" = c(0, 0, 0, 0, 0, 0, 0, –1, 1))
> Library(multcomp)
> compsaost1p <- glht(AOST1P.2, contrast.matrix)
> summary(compsaost1p)
```

(*table continues*)

Appendix C (p. 2 of 2)

R Code Used in the Analyses

Mixed-effects binomial logistic regression

Model building

```
> aos0.glm = glm(Accuracy ~ 1, family = binomial(link = "logit"), data = aospedata)
> aos0.glmer = glmer(Accuracy ~ (1|ID), data = aospedata, family = binomial(link = "logit"))
```

Testing the random effect

```
> aic.aosglmer <- AIC(logLik(aos0.glmer))
> aic.aosglm <- AIC(logLik(aos0.glm))
> aic.aosglmer; aic.aosglm
```

Model fitting

```
> aosglmer.glmulti <- function(formula, data, random="", . . .){glmer(paste(deparse(formula), random), family = binomial, data =
aospedata, control = glmerControl(optimizer = "bobyqa"), . . .)}
> form_glmulti = as.formula(paste("Accuracy ~ Syllable + Lexical + Position"))
> aosmfit <- glmulti(form_glmulti,random="+(1 | ID)", data = aospedata, method = "h", fitfunc = aosglmer.glmulti, crit = "bic", intercept =
TRUE, marginality = FALSE, level = 2)
```

Extract the best models

```
< top <- weightable(aosmfit)
< top <- top[1:5,]
< top
```

Inspect and define final minimal model

```
< aos.glmer <- glmer(Accuracy ~ (1|ID) + Syllable + Lexical + Position, family = binomial, control = glmerControl(optimizer = "bobyqa"),
data = aospedata)
< summary(aos.glmer, corr = F)
```

Test whether the final minimal model performs significantly better than the minimal base-line model

```
< sigfit <- anova(aos.glmer, aos0.glmer, test = "Chi")
< sigfit
```

Extracting model fit parameters

```
< probs = 1/(1+exp(-fitted(aos.glmer)))
< probs = binomial()$linkinv(fitted(aos.glmer))
< somers2(probs, as.numeric(aospedata$Accuracy))
```

Model Diagnostics

```
< sjPlot::tab_model(aos.glmer)
```

Note. The bold and underscore codes indicated the specific parameters for the analysis.

Appendix D (p. 1 of 2)

The Model Fit for Normalized f_o Values by the Effects of Group, Syllable Structure, Lexical Status, and Syllable Position

Effects of group		*df*	χ^2				
High-level tone (T1)	Intercept	2	8.986*				
	Linear term	2	2.816				
	Quadratic term	2	8.050				
High-rising tone (T2)	Intercept	2	4.055				
	Linear term	2	14.771***				
	Quadratic term	2	9.064*				
Low-falling tone (T4)	Intercept	2	0.627				
	Linear term	2	1.946				
	Quadratic term	2	19.781***				
Effects of syllable structure (V vs. CV structure)		**AOS**		**nAOS**		**HC**	
		df	χ^2	*df*	χ^2	*df*	χ^2
High-level tone (T1)	Intercept	1	95.6045***	1	144.2412***	1	216.9947***
	Linear term	1	8.1801**	1	0.1143	1	10.5215**
	Quadratic term	1	1.6314	1	3.1017	1	6.3507*
High-rising tone (T2)	Intercept	1	1.3842	1	455.5864***	1	1467.643***
	Linear term	1	8.8732**	1	6.3864*	1	84.333***
	Quadratic term	1	5.3136*	1	10.2428**	1	22.366***
Low-falling tone (T4)	Intercept	1	0.2854	1	144.1886***	1	413.9891***
	Linear term	1	0.3933	1	2.8542	1	58.7534***
	Quadratic term	1	1.7163	1	1.7801	1	7.2448**
Effects of lexical status (words vs. nonwords)		**AOS**		**nAOS**		**HC**	
		df	χ^2	*df*	χ^2	*df*	χ^2
High-level tone (T1)	Intercept	1	1.6217		41.8395***		34.5737***
	Linear term	1	1.1352		2.9140		2.3859
	Quadratic term	1	1.1192		0.1684		0.9603
High-rising tone (T2)	Intercept	1	0.9766		10.3936**		15.6822***
	Linear term	1	10.3914**		2.0854		0.2718
	Quadratic term	1	0.0702		0.0718		0.8727
Low-falling tone (T4)	Intercept	1	21.8702***		40.3592***		20.4802***
	Linear term	1	8.2287**		0.2447		0.0154
	Quadratic term	1	0.0000		0.0524		0.3496

(*table continues*)

Appendix D (p. 2 of 2)

The Model Fit for Normalized f_o Values by the Effects of Group, Syllable Structure, Lexical Status, and Syllable Position

Effects of group		*df*	χ^2				
Effects of syllable position (initial vs. medial vs. final)		**AOS**		**nAOS**		**HC**	
		df	χ^2	*df*	χ^2	*df*	χ^2
High-level tone (T1)	Intercept	2	316.3741***	2	564.5615***	2	618.851***
	Linear term	2	179.5028***	2	96.8662***	2	269.345***
	Quadratic term	2	6.4173*	2	1.5048	2	19.984***
High-rising tone (T2)	Intercept	2	1968.859***	2	704.506***	2	1545.58***
	Linear term	2	90.938***	2	364.736***	2	773.12***
	Quadratic term	2	26.912***	2	29.365***	2	187.45***
Low-falling tone (T4)	Intercept	2	2103.635***	2	107.205***	2	634.26***
	Linear term	2	486.914***	2	146.668***	2	342.09***
	Quadratic term	2	30.007***	2	17.842***	2	67.48***

Note. V = vowel structure; CV = consonant–vowel structure; AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls.

*$p < .05$ **$p < .01$ ***$p < .001$.

Appendix E (p. 1 of 3)

Statistical Data From the Growth Curve Analysis of the Normalized Fundamental Frequency Values in Tone 1, Tone 2, and Tone 4 Syllables for the Effects of Group (E1), Syllable Structure and Lexical Status (E2), and Syllable Position (E3)

Table E1. Parameter estimates for the analysis of the effects of group on normalized fundamental frequency values in Tone 1, Tone 2, and Tone 4 syllables.

Variable		Estimate	*SE*	*z*
High-level tone (T1)				
AOS vs. nAOS	Intercept	0.3867	0.1261	3.065*
	Linear	0.1869	0.1141	1.634
	Quadratic	−0.1226	0.1064	−1.153
AOS vs. HC	Intercept	0.6028	0.1261	4.779***
	Linear	0.1782	0.1141	1.558
	Quadratic	−0.3350	0.1063	−3.149*
nAOS vs. HC	Intercept	−0.2161	0.1261	−1.713
	Linear	0.0087	0.1141	0.076
	Quadratic	0.2124	0.1064	1.996
High-rising tone (T2)				
AOS vs. nAOS	Intercept	−0.1183	0.0728	−1.624
	Linear	0.5672	0.1417	4.004***
	Quadratic	0.2335	0.0730	3.200*
AOS vs. HC	Intercept	−0.1540	0.0728	−2.115
	Linear	0.7720	0.1417	5.449***
	Quadratic	0.1951	0.0730	2.675
nAOS vs. HC	Intercept	0.03574	0.0728	0.491
	Linear	−0.2048	0.1417	−1.445
	Quadratic	0.0384	0.0730	0.526
Low-falling tone (T4)				
AOS vs. nAOS	Intercept	−0.2702	0.1074	−2.517
	Linear	−0.0345	0.1094	−0.315
	Quadratic	0.2530	0.0535	4.733***
AOS vs. HC	Intercept	−0.4417	0.1074	−4.114***
	Linear	−0.2081	0.1094	−1.903
	Quadratic	0.2974	0.0535	5.563***
nAOS vs. HC	Intercept	0.1715	0.1074	1.597
	Linear	0.1736	0.1094	1.588
	Quadratic	−0.0444	0.0535	−0.830

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls.

*$p < .05$. ***$p < .001$.

Appendix E (p. 2 of 3)

Statistical Data From the Growth Curve Analysis of the Normalized Fundamental Frequency Values in Tone 1, Tone 2, and Tone 4 Syllables for the Effects of Group (E1), Syllable Structure and Lexical Status (E2), and Syllable Position (E3)

Table E2. Parameter estimates for the analysis of effects of syllable structure and lexical status on normalized fundamental frequency values in Tone 1, Tone 2, and Tone 4 syllables.

Variable		*t* values		
		AOS	**nAOS**	**HC**
Syllable structure (V vs. CV structure)				
High-level tone (T1)	Intercept	–9.7895***	12.0289***	14.7693***
	Linear	2.8604**	0.3381	–3.2445**
	Quadratic	–1.2773	–1.7612	–2.5202*
High-rising tone (T2)	Intercept	–1.1769	21.4537***	38.9901***
	Linear	2.9794**	–2.5278*	9.1957***
	Quadratic	–2.3052*	–3.2008**	–4.7303***
Low-falling tone (T4)	Intercept	–0.5340	12.0271***	20.4630***
	Linear	0.6271	1.6895	7.6709***
	Quadratic	–1.3101	1.3342	–2.6918**
Lexical status (word vs. nonword items)				
High-level tone (T1)	Intercept	1.6289	–6.4813***	–5.8901***
	Linear	–1.0656	–1.7073	1.5449
	Quadratic	–1.0580	0.4104	0.9800
High-rising tone (T2)	Intercept	–0.9891	–3.2259**	–3.9601***
	Linear	3.2250**	1.1112	–0.5214
	Quadratic	–0.2650	0.2679	0.9342
Low-falling tone (T4)	Intercept	–4.6841***	–6.3638***	–4.5295***
	Linear	2.8696**	0.4947	–0.1240
	Quadratic	0.0067	0.2288	0.5913

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls; V = vowel structure; CV = consonant–vowel structure.

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Appendix E (p. 3 of 3)

Statistical Data From the Growth Curve Analysis of the Normalized Fundamental Frequency Values in Tone 1, Tone 2, and Tone 4 Syllables for the Effects of Group (E1), Syllable Structure and Lexical Status (E2), and Syllable Position (E3)

Table E3. Parameter estimates for the analysis of effect of syllable position on normalized fundamental frequency values in Tone 1, Tone 2, and Tone 4 syllables.

Variable		AOS	nAOS	HC
High-level tone (T1)				
Initial vs. medial	Intercept	1.101	–21.155***	–25.175***
	Linear	–0.314	3.675**	12.395***
	Quadratic	–1.387	–1.113	–1.477
Initial vs. final	Intercept	–14.937***	–20.285***	–12.580***
	Linear	–11.779***	9.754***	15.580***
	Quadratic	1.142	–0.110	2.916*
Medial vs. final	Intercept	–16.038***	0.869	12.594***
	Linear	–11.465***	6.078***	3.185*
	Quadratic	2.529	1.003	4.393***
High-level tone (T2)				
Initial vs. medial	Intercept	31.178***	1.289	9.975***
	Linear	–2.715	0.423	–3.029*
	Quadratic	–2.096	–1.575	–2.578
Initial vs. final	Intercept	–12.943***	–22.659***	–29.179***
	Linear	–9.287***	16.816***	22.700***
	Quadratic	3.063*	3.704**	10.379***
Medial vs. final	Intercept	–44.121***	–23.948***	–39.154***
	Linear	–6.572***	16.397***	25.729***
	Quadratic	5.159***	5.279***	12.956***
Low-falling tone (T4)				
Initial vs. medial	Intercept	9.928***	–9.487***	6.865***
	Linear	–1.350	–11.928***	–14.870***
	Quadratic	–1.163	2.801*	2.754*
Initial vs. final	Intercept	–35.107***	–8.432***	–17.881***
	Linear	–19.853***	–4.043***	–17.088***
	Quadratic	4.056***	4.139***	8.085***
Medial vs. final	Intercept	–45.035***	1.055	–24.747***
	Linear	–18.503***	7.885***	–2.218
	Quadratic	5.219***	1.338	5.330***

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech post-stroke; HC = healthy controls.

*$p < .05$. **$p < .01$. ***$p < .001$.

Appendix F

Descriptive Statistics for the Perceptual Ratings of Tone Accuracy of the Tone 1, Tone 2, and Tone 4 Syllables in the Three Groups of Participants

Variable		M (SD)		
		AOS	nAOS	HC
Tone				
High level (T1)		0.72 (0.451)	0.89 (0.307)	1.00 (0.019)
High rising (T2)		0.26 (0.437)	0.75 (0.431)	0.93 (0.207)
Low falling (T4)		0.55 (0.498)	0.86 (0.346)	1.00 (0.000)
Syllable structure				
High level (T1)	V	0.67 (0.469)	0.90 (0.299)	1.00 (0.000)
	CV	0.70 (0.457)	0.86 (0.349)	1.00 (0.029)
High rising (T2)	V	0.20 (0.403)	0.69 (0.464)	0.94 (0.234)
	CV	0.24 (0.428)	0.77 (0.424)	0.96 (0.200)
Low falling (T4)	V	0.51 (0.500)	0.84 (0.363)	1.00 (0.000)
	CV	0.56 (0.499)	0.85 (0.361)	1.00 (0.000)
Lexical status				
High level (T1)	W	0.95 (0.218)	0.98 (0.140)	1.00 (0.000)
	NW	0.69 (0.464)	0.88 (0.320)	1.00 (0.020)
High rising (T2)	W	0.62 (0.487)	0.94 (0.161)	0.99 (0.115)
	NW	0.21 (0.410)	0.73 (0.445)	0.95 (0.215)
Low falling (T4)	W	0.77 (0.422)	0.97 (0.161)	1.00 (0.000)
	NW	0.52 (0.500)	0.85 (0.360)	1.00 (0.000)
Syllable position				
High level (T1)	Initial	0.67 (0.472)	0.96 (0.200)	1.00 (0.035)
	Medial	0.69 (0.462)	0.88 (0.328)	1.00 (0.000)
	Final	0.71 (0.453)	0.83 (0.380)	1.00 (0.000)
High rising (T2)	Initial	0.15 (0.356)	0.62 (0.486)	0.91 (0.287)
	Medial	0.22 (0.415)	0.70 (0.460)	0.94 (0.232)
	Final	0.29 (0.454)	0.88 (0.325)	1.00 (0.000)
Low falling (T4)	Initial	0.48 (0.500)	0.87 (0.334)	1.00 (0.000)
	Medial	0.38 (0.485)	0.86 (0.349)	1.00 (0.000)
	Final	0.71 (0.454)	0.82 (0.381)	1.00 (0.000)

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls; V = vowel structure; CV = consonant–vowel structure; W = word items; NW = nonword items.

Appendix G

Descriptive Statistics for Repetition Duration of All Items and Items With Different Syllable Structures and Lexical Status in the Three Groups of Participants

Variable	*M* (*SD*)		
	AOS	nAOS	HC
Repetition duration (s)			
First set	1.43 (0.64)	1.48 (0.58)	1.08 (0.13)
Second set	4.75 (1.51)	5.04 (2.29)	3.20 (0.56)
Third set	5.24 (3.54)	4.70 (2.29)	2.79 (0.33)
Overall	4.07 (2.24)	4.21 (2.51)	2.68 (0.99)
Syllable structure–V			
First set	1.53 (0.60)	1.44 (0.48)	1.16 (0.15)
Second set	4.83 (1.48)	5.54 (2.72)	3.37 (0.60)
Overall	4.00 (1.94)	4.54 (2.94)	2.81 (1.10)
Syllable structure–CV			
First set	1.33 (0.68)	1.32 (0.68)	1.04 (0.11)
Second set	4.68 (1.55)	4.54 (1.61)	3.03 (0.46)
Overall	3.84 (2.00)	3.77 (1.97)	2.52 (0.96)
Lexical status–words			
First set	1.43 (0.63)	1.45 (0.71)	1.12 (0.14)
Third set	5.04 (2.85)	4.51 (1.96)	2.76 (0.35)
Overall	2.61 (2.40)	2.51 (1.88)	1.66 (0.82)
Lexical status–nonwords			
First set	1.43 (0.66)	1.31 (0.43)	1.08 (0.15)
Third set	5.44 (4.13)	4.89 (2.59)	2.83 (0.31)
Overall	2.79 (3.10)	2.57 (2.24)	1.64 (0.86)

Note. AOS = adult speakers with apraxia of speech poststroke; nAOS = adult speakers without apraxia of speech poststroke; HC = healthy controls; V = vowel structure; CV = consonant–vowel structure.