Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: www.tandfonline.com/journals/paph20

Tone-vowel interaction and co-articulation in Cantonese speakers with apraxia of speech and co-existing aphasia: a preliminary study

Yixin Zhang, Eddy C.H. Wong & Min Ney Wong

To cite this article: Yixin Zhang, Eddy C.H. Wong & Min Ney Wong (27 Dec 2024): Tone-vowel interaction and co-articulation in Cantonese speakers with apraxia of speech and co-existing aphasia: a preliminary study, Aphasiology, DOI: 10.1080/02687038.2024.2441203

To link to this article: https://doi.org/10.1080/02687038.2024.2441203

APHASIOLOGY
https://doi.org/10.1080/02687038.2024.2441203

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Tone-vowel interaction and co-articulation in Cantonese speakers with apraxia of speech and co-existing aphasia: a preliminary study

Yixin Zhang [a*], Eddy C.H. Wong [a,b*] and Min Ney Wong [a,c,d,e]

[a]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China; [b]Unit of Human Communication, Language, and Development, Faculty of Education, The University of Hong Kong, Hong Kong, China; [c]Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China; [d]Research Institute for Smart Ageing, The Hong Kong Polytechnic University, Hong Kong, China; [e]The HK PolyU-PKU Research Centre on Chinese Linguistics, The Hong Kong Polytechnic University, Hong Kong, China

ABSTRACT

Background: Speakers with apraxia of speech (AOS) usually produce segmental and prosodic errors that influence their speech intelligibility. Literature on AOS in tonal language speakers is sparse compared to that in non-tonal language. Also, the existing research often made static, isolated analyses, leaving the production and co-articulation between segments and suprasegmental entities in tonal languages under-investigated.

Aims: This preliminary study aims to fulfill the aforementioned research gaps by investigating vowel-tone interaction and tonal and vocalic co-articulation in Cantonese post-stroke speakers with AOS.

Methods: Five Cantonese adults with AOS post-stroke, five adults without AOS post-stroke, and five healthy controls performed the Tone Sequencing Task (TST), a task adapted from oral diadochokinetic tasks that required five rapid repetitions of 3-syllable items formed by three different Cantonese vowels and three different Cantonese tones. The quality of vowels was indexed by midpoint formant values and euclidean distances between the vowels. Within-speaker variation was assessed by coefficients of variance. Co-articulation was indexed by onset and offset formant or f_0 values. The effects of the participant groups, the positions of tone-syllable in the TST stimuli, and the tones carried by vowels/carrying vowels were evaluated with linear mixed effect models.

Results: Cantonese-speaking adults with AOS had difficulty in producing distinctive vowels and tones. They also showed large within-speaker variation in vowel production but reduced tone contrast, especially at the final positions. The disrupted anticipatory co-articulation between vowels as well as between tones further suggested that the speakers with AOS could not sequence segments and suprasegments simultaneously.

ARTICLE HISTORY
Received 6 September 2023
Accepted 8 December 2024

KEYWORDS
Apraxia of speech; co-articulation; tone-vowel interaction; tonal language; Cantonese

CONTACT Yixin Zhang yixin-yz.zhang@polyu.edu.hk Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong HJ613, China

*These two authors contribute equally to the present study. Therefore, they are co-first authors of the article.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02687038.2024.2441203.

1. Introduction

Apraxia of speech (AOS) is a type of neurological speech disorder that can be acquired in adulthood following a brain injury (e.g., stroke) or neurodegenerative disease (Duffy, 2020). Stroke is the most common cause of AOS (Duffy, 2020) and AOS commonly co-occurs with aphasia and/or dysarthria. Speakers with AOS have an impairment of motor planning of speech movement (van der Merwe, 2021), resulting in speech sounds and prosodic errors that influence speech intelligibility.

Some perceptually-related clinical features of AOS overlap with aphasia and dysarthria, making these features non-discriminative. For example, inconsistent variants of sound errors can be found in speakers with aphasia with co-existing AOS as well as in speakers with aphasia without co-existing AOS but with phonemic paraphasia (Duffy, 2020; K. L. Haley et al., 2021). In contrast, some clinical features are considered discriminative, and are summarised in clinical tools such as the Apraxia of Speech Rating Scale (ASRS; Strand et al., 2014).

1.1. Segmental performances

Segmental errors produced by speakers with AOS include sound distortions, distorted sound substitutions (i.e., phoneme errors that are phonetically distorted), and distorted sound or syllable additions, such as intrusive schwa (Duffy, 2020; Duffy et al., 2023; K. L. Haley et al., 2017). With regard to vowel production in particular, the existing studies have reported mixed results. While some studies showed that speakers with AOS produced vowels different from healthy controls, others reported comparable vowel production between speakers with AOS and healthy controls. Kent and Rosenbek's acoustic analysis (1983) of vowels produced by speakers with AOS in connected speech (i.e., conversations, picture descriptions, and readings) was more consistent with the early phonetic transcriptions and found a general imprecision of articulatory positioning of vowels (measured by acoustic parameters including vocalic formants, duration, and intensity). The authors suggested that listeners may have perceived small acoustic deviations as distortions and large acoustic deviations as categorical substitutions. Similarly, some other research also suggested that some speakers with AOS and co-existing aphasia produced vowels with formant patterns that deviated from healthy controls, but the degree and nature of deviation varied across individuals, with trials showing large inter- and within-speaker variations (K. Haley, 2004; K. L. Haley et al., 2000, 2001). Furthermore, some studies also found that the height dimension (acoustic parameter, F1) of vowels is disrupted more often than frontness (acoustic parameter, F2) among speakers with AOS (K. Haley, 2004; K. L. Haley et al., 2000, 2001; Odell et al., 1991). By contrast, in Jacks et al. (2010), the vowels in six frequently occurring English words (i.e., hid, head, hat, hot, hub, hoot) produced by seven adults with AOS and co-existing aphasia did not differ from those produced by the healthy controls, in terms of absolute formant values, inter-vowel distance, and within-speaker trial-to-trial formant variability.

1.2. Prosodic performances

With regard to prosody, speakers with AOS who speak non-tonal languages with lexical stress are characterised by slow speech rates, lengthened co-articulatory transitions, reduced intensity variation across syllables, and inappropriate pause between sounds and syllables (Duffy, 2020; McNeil et al., 2017; Strand et al., 2014; Utianski et al., 2018). With regard to word stress production which involves simultaneous variations of pitch, loudness, and duration, speakers with AOS have been found to have difficulty in varying duration, while pitch and intensity produced by them are similar to those with aphasia only and healthy controls (K. L. Haley & Jacks, 2019; Vergis et al., 2014). However, on utterance-level, the disfluent articulation corrupts the regular rhythm and melody of speech. Difficulty of varying pitch was found in sentences produced by speakers with AOS. To be specific, compared to speakers without AOS, they have been found to produce utterances with reduced fundamental frequency (f_0) contour as well as persistent terminal f_0 falling (Boutsen & Christman, 2002; Kent & Rosenbek, 1983). The inconsistent findings may be attributed to the differences in pitch and loudness requirements between words and longer utterances. Specifically, single-word tasks in English may not necessitate distinct pitch and loudness contours, whereas phrases or sentences may.

1.3. Tone production in AOS

Although pitch variational measures were not promising in differentiating AOS in English speakers, that does not mean they are also minimally useful in other languages, such as tonal languages, where f_0 is the primary cue to differentiate the meaning of words. Pitch variations were investigated in tonal language speakers with AOS or childhood apraxia of speech (a similar speech disorder with an onset in childhood; W. Chen et al., 2022; Wong, Wong, Chen, et al., 2024; Wong, Wong, & Velleman, 2024; Wong et al., 2021). W. Chen et al. (2022) study on tone and vowel productions in Mandarin speakers with AOS post-stroke reported that speakers with AOS had equally frequent errors in tones and vowels when producing monosyllabic words; however, larger acoustic deviations were found in tonal targets than in vowel targets. In perceptual analysis, native Mandarin-speaking listeners perceived deviated tones as more “correct” than deviated vowels. The authors discussed the possibility that such asymmetric tolerance may be due to the simple Mandarin tonal system (i.e., one level and three contour tones carried by fully-stressed syllables plus a neutral tone carried by unstressed syllables), compared to other tonal languages or dialects (e.g., Cantonese and Min Dialect). Wong, Wong, Chen, et al. (2024) investigated tone production in Cantonese, a tonal language with six contrastive tones: high-level tone (T1), high-rising tone (T2), mid-level tone (T3), low-falling tone (T4), low-rising tone (T5), and low-level tone (T6) (Figure 1). They reported degraded pitch-variations skills in Cantonese speakers with AOS post-stroke compared to healthy controls and those with aphasia and/or dysarthria but no AOS post-stroke. In particular, Cantonese speakers with AOS post-stroke have more difficulty in contour tones with a rising than falling pitch, which was not observed in the control groups. In fact, Sundberg (1979) argued that a falling f_0 contour is physiologically less demanding than a rising contour, so that populations who are more sensitive to physiological constraints, like young children (e.g., Ballard et al., 2012; Snow, 2007), and speakers with AOS (Wong, Wong, Chen, et al., 2024) may show preference to falling contours when realising lexical tones or lexical stress.

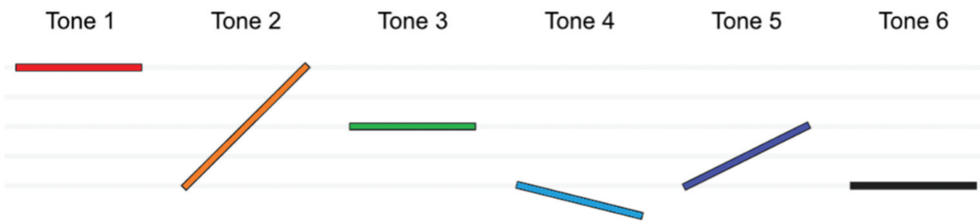

Figure 1. Illustrative sketch of the f_0 contours of the six Cantonese tones on the stave.

1.4. Tone-vowel interaction

Although W. Chen et al. (2022) and Wong, Wong, Chen, et al. (2024) investigated pitch performance, they did not investigate the interaction between vowels and tones, nor the co-articulation between vowels and tones. However, tone production is suggested as within a word's motor programme instead of a prosodic feature in the Directions Into Velocities of Articulators (DIVA) model (Guenther, 2016), so that tones and vowels may either compete or cooperate with each other. Therefore, an investigation of vowels and tones and their interaction may add to the literature about the role of tone in the speech processing model.

In addition, the tones carried by vowels as well as the neighbouring vowels influence the articulatory positioning in vowel production and hence the formant patterns of vowels, but there was still a lack of report of the tone-vowel interaction and co-articulation in tone-language speakers with AOS. It is well established that tone production involves laryngeal movement; namely, a tone with high f_0 requires a raised larynx position, which further shortens a person's vocal tract between the glottis and lips (see Ohala, 1973; Yip, 2002a). As a result, tones are found to influence the articulatory positioning in vowel production (Erickson et al., 2004; Hoole & Hu, 2004) and hence the formant patterns of vowels (Torng, 2000; Torng et al., 2000; Zee, 1980). Specifically, several studies reported that formants of vowels tended to be raised by high tonal targets, with F1 more susceptible to tonal impacts than F2 (Hombert et al., 1979; Hoole & Hu, 2004; Zee, 1980). Research suggests that healthy speakers may compensate for the first two formants to retain the phonemic identity of vowels when there are great changes in f_0 (Ainsworth, 1975; Parmenter et al., 1933; Torng, 2000; Zee, 1980). However, speakers with AOS are unable to adjust the acoustic variables, resulting in more disrupted F1 than F2 (K. Haley, 2004; K. L. Haley et al., 2000, 2001; Odell et al., 1991). It is reasonable to predict that tonal language speakers with AOS may have a great difficulty in vowel productions, given that they need to maintain the intelligibility of tone and vowels at the same time.

1.5. Co-articulation

In addition to the interaction between the tones and the carrying vowels, the neighboring phonemes and tonemes also induce changes in articulation and acoustic signals, which is referred to as co-articulation. Co-articulation can result from the activation of a single articulatory gesture or the interaction between articulatory gestures. It can be divided into anticipatory (i.e., preceding phonemes or tonemes influenced by the following ones) or carryover (i.e., following phonemes or tonemes influenced by the preceding ones). A commonly accepted view regarding co-articulation is that carryover effects are more dependent on mechanoinertial factors than anticipatory effects, while anticipatory effects reflect more speech planning (Lindblom & Studdert-Kennedy, 1967; Recasens, 1984; Whalen, 1990). This view has been empirically supported. For instance, research found that people suffering from motor planning difficulty such as AOS showed more problems in anticipatory co-articulation than in carryover co-articulation (Ziegler & von Cramon, 1986). Furthermore, one of the discriminative features in speakers with AOS regarding co-articulatory skills is syllable segregation (Duffy, 2020), which is caused by reduced anticipatory co-articulation. Ziegler and von Cramon (1986) found that their German-speaking patient with AOS lacked co-articulatory cohesion in speech, in particular a consistent delay in the initiation of anticipatory vowel gestures, resulting in syllable segregation. Southwood et al. (1997) also found that the anticipatory co-articulatory gestures of the English-speaking patient with AOS were delayed or distorted, regardless of words or speaking rates. It is worth noting, however, that saying that the carryover effects are more dependent on "mechanoinertial factors" does not deny that it may still involve some speech motor planning, namely, it is possible that planning of the following phoneme needs accommodation of the position or state of articulators from the preceding gesture. For instance, Ostry et al. (1996) found that both anticipatory and carryover jaw movements correlated to the amplitude and duration of the preceding and following gestures, and they were influenced by both central control and muscle mechanics. Nevertheless, the widely reported asymmetry between anticipatory and carryover effects (i.e., carryover effects are usually stronger) may still suggest that, in addition to central control, mechanoinertial factors may exert a larger influence on carryover effects.

Vowel-to-vowel coarticulation refers to the coarticulatory effects of one vowel on another across one or more intervening consonants. Acoustically, vowel-to-vowel co-articulation is manifested as the assimilation of the F1 and F2 values of a vowel to its neighboring vowels due to overlap between lip and tongue gestures, and the co-articulatory effects are usually stronger at vocalic edges than midpoints across various languages (Magen, 1997; S. Y. Manuel & Krakow, 1984; P. P. Mok, 2012, 2013; Recasens, 1984). In other words, a vowel preceding or following a high front vowel tends to have a higher offset or onset, namely, an offset or onset produced with a more closed mouth, a higher and more front tongue body, as well as a smaller F1 and a larger F2. The degree of vowel-to-vowel coarticulation varies across languages, and the existence of secondary articulation and the phonemic inventory size of vowels have been used to explain such variations (e.g., Choi & Keating, 1991; S. Manuel, 1999). However, P. P. Mok (2013) found that the relatively crowded phonemic vowel space did not reduce the strength of vowel-to-vowel co-articulation in Cantonese when compared to Mandarin, and she reported stronger carryover effects in both F1 and F2 of [i], [a], and [u] than anticipatory effects.

Given that vowel performance can be influenced by tone, as mentioned, another important aspect that requires investigation in Cantonese speakers with AOS is tone-to-tone co-articulation. Tones produced in sequences also tend to adapt to their adjacent tones at edges, and there are usually greater co-articulatory effects between antagonistic tones than compatible tones. Tonal co-articulation differs from segmental co-articulation in several ways. Firstly, tonal co-articulation is usually restricted to contiguous tones (Gu & Lee, 2009; B. Li et al., 2020; Shen, 1990), whereas vowel-to-vowel co-articulation may extend from one vowel to another across intervening consonants or even medial schwa in unstressed syllables (Magen, 1997; S. Y. Manuel & Krakow, 1984; P. P. Mok, 2013; P. K. P. Mok & Hawkins, 2004; Recasens, 1984). In many languages, carryover co-articulation between tones tends to be more prominent than anticipatory effects (see Xu, 1997 for Mandarin; see Han & Kim, 1974 for Vietnamese; see Potisuk et al., 1997 for Thai), though in some dialects like Malaysian Hokkein (Chang & Hsieh, 2012) and Nanjing Mandarin (S. Chen et al., 2018), anticipatory co-articulation is of comparable applicability and strength to carryover co-articulation. In Cantonese, B. Li et al. (2020) found that the strength of carryover co-articulation depends on the type of preceding tone. For instance, the carryover co-articulation triggered by the low tone carried throughout the entire tonal contour, but the co-articulation triggered by the high tone did not. In addition, although anticipatory co-articulation between most tones is assimilatory like that between segments, a following low tone tends to raise its preceding tone, resulting in "pre-low raising" (see Xu, 1997 for Mandarin; see J. Gandour et al., 1996; A. Lee et al., 2021 for Thai; see Gu & Lee, 2009; A. Lee et al., 2021 for Cantonese). Moreover, such anticipatory co-articulation between tones, though often less prominent, can be more fixed and extended to the onset of the whole sonorant part of the preceding syllables rather than affecting the edges only (Brunelle, 2009; Gu & Lee, 2009; Shen, 1990). Perceptually, native Cantonese speakers are also found to use anticipatory rising but not pre-low raising in pitch contours to resolve lexical competition (Qin & Zhang, 2022); nevertheless, such findings still indicate that tonal co-articulation plays a relatively important role in Cantonese speech communication.

1.6. Purposes and hypotheses

The present study aims to fill in the aforementioned gaps by investigating the tone-vowel interaction, vowel-to-vowel, and tone-to-tone co-articulation in Cantonese speakers with AOS and co-existing aphasia post-stroke in comparison with matched patients without AOS but with aphasia post-stroke, and healthy adults. Based on the previous studies, we made the following hypotheses:

H1. Producing vowels with tones in sequences will be more difficult for speakers with AOS, particularly for vowel height (F1).

H2. Speakers with AOS will have disrupted anticipatory co-articulation when compared to the two control groups; to be specific, the offset of the preceding vowels or tones would hardly be similarized to the following ones in height and frontness in the AOS group.

H3. Speakers with AOS will not have disrupted carryover co-articulation, relative to the two control groups; namely, the onset of the following vowels or tones would be similarized to the preceding ones in height and frontness in all three groups.

2. Methodology

This study was approved by the PolyU Institutional Review Board (reference number: HSEARS20210330007).

2.1. Participants

Five adults with AOS post-stroke (AOS) and co-existing aphasia, five adults without AOS but with aphasia post-stroke (nAOS) and five neurologically typical participants (HC) participated in the study. Five of the participants (three from the AOS and two from the nAOS groups) also had mild co-existing dysarthria. The control groups (nAOS and HC) were age-matched with the participants in the AOS group. Participants in the AOS and nAOS groups were mainly recruited from the university speech-language pathology clinic while the participants in the HC groups were recruited from the caregivers or family members of the participants, as well as open recruitment within the university. All the participants provided written consent before joining the study. The clinical and demographic information of the participants is summarised in Table 1.

2.2. Procedure

All the participants received a three-hour speech and language assessment by two master speech-language pathology students under supervision of a registered speech-language pathologist (SLP) (author: EW). In particular, the two student clinicians provided assessment to 11 (4 in the AOS group, 5 in the nAOS group, and 2 in the HC group) and 4 participants (1 in the AOS group and 3 in the HC group), respectively. The assessment protocol included case history taking, speech sample collection, the Cantonese Aphasia Battery (CAB) (Yiu, 1992), the Computerized Revised Token Test-Cantonese (CRTT-Cantonese) (Bakhtiar et al., 2020), the unofficial translated Cantonese version of the Frenchay Dysarthria Assessment-2 (FDA-2) (Enderby & Palmer, 2008, translated by SLP clinical educators who work at a university with no report of its psychometric properties), and the Cantonese Tone Identification Test (K. Y. S. Lee, 2012). The sessions were video-recorded. The diagnosis of aphasia was confirmed if the participant fell below the cut-off point for CAB Aphasia Quotient (<96.4).

Two experienced SLPs who were independent to the study and blinded to the participants' speech and language diagnoses participated in this study to confirm AOS diagnoses in the participants. The SLPs reviewed the video-recorded assessment sessions and completed the Apraxia of Speech Rating Scale (ASRS) 3.0 (Utianski et al., 2018) independently before the online meeting, which aimed to produce a single consensus rating to be used in all further analysis. The assessment results of the participant are presented in Table 1.

The dysarthria diagnoses were confirmed by the author EW, through reviewing the videos of the assessment sessions, in particular, the performances on the unofficial translated version of the FDA-2. Though mild dysarthria was reported in three participants with AOS post-stroke, the results of the laryngeal subtest of the FDA-2 showed that these participants had no abnormality in pitch gliding (i.e., sing a scale with at least six notes). Their difficulties in laryngeal functions included reduced sustained phonation, limited loudness variation, and atypical voice quality.

All participants performed the Tone Sequencing Task (TST) in the same day after the assessment session. The TST was adopted from the oral diadochokinetic (DDK) task, which requires participants to repeat items as fast as they could. The participants in this study were asked to repeat the items once as baseline and then five times even when their

Table 1. Demographic information of the participants.

Group	ID	Gender	Age	Months post-onset	CAB AQ	Aphasia type	CRTT-L-Cantonese score	CRTT-R-$_{WF}$-Cantonese score	CANTIT percentile	Dysarthria type	Dysarthria severity	AOS severity
AOS	#1	M	52	94	76	Conduction	NA	NA	<4	UUMN	Mild	Mild
	#2*	M	50	84	96.2	Anomic	NA	NA	16	Flaccid & Hypokinetic	Mild	Mild
	#5	M	56	165	82.3	Anomic	10.36	10.76	16	Flaccid	Mild	Mild
	#10	M	43	61	47	Broca's	10.55	7.48	<4	NA	NA	Moderate
	#17	F	58	19	80.8	Anomic	11.90	12.74	<4	NA	NA	Mild
nAOS	#3	M	53	31	76.2	Transcortical motor	NA	NA	44	Flaccid	Mild	NA
	#4	F	55	63	72.6	Transcortical sensory	9.88	9.39	16	Flaccid	Mild	NA
	#7	F	49	58	95.8	Anomic	12.36	12.80	44	NA	NA	NA
	#9	M	59	52	91.2	Anomic	12.74	12.94	82	NA	NA	NA
	#11	M	64	19	77	Anomic	NA	NA	44	NA	NA	NA
HC	#13	M	59	NA	99	NA	NA	NA	44	NA	NA	NA
	#15	F	54	NA	100	NA	NA	NA	44	NA	NA	NA
	#19	M	45	NA	100	NA	14.5	14.65	82	NA	NA	NA
	#20	M	50	NA	99.8	NA	14.0	14.29	44	NA	NA	NA
	#21	M	51	NA	100	NA	14.2	14.35	82	NA	NA	NA

Abbreviations. AOS = apraxia of speech group; nAOS = no apraxia of speech; HC = Healthy control; CAB = Cantonese Aphasia Battery; AQ = Aphasia Quotient; UMN = Unilateral upper motor neuron dysarthria; Flaccid = Flaccid dysarthria; Hypokinetic = Hypokinetic dysarthria; CRTT-L-Cantonese = Computerized Revised Token Test – Listening (Cantonese version); CRTT-R-$_{WF}$-Cantonese = Computerized Revised Token Test – Reading-Word Fade (Cantonese version); CANTIT = The Cantonese Tone Identification Test; WNL = within normal limits; ASRS = Apraxia of Speech Rating Scale; NA = not applicable. *Note*. * Participant 2 had encephalitis one month after the stroke.

baseline performances were incorrect. The TST was modified based on Wong et al. (2021). There were 108 items, which were formed with early acquired initial consonants (i.e., [m, j, t, n]) and vowels (i.e., [a, ɔ, i, ɛ]), as well as three early acquired Cantonese tones, including high-level tone (T1), high-rising tone (T2), and low-falling tone (T4). T1, T2, and T4 are chosen among the six Cantonese tones because they involve the upper and lower limits of tone (i.e., the high and low tones) and all three tone contours in Cantonese (i.e., level, rising, and falling). The vowels are chosen because they represent different vowel height and frontness in Cantonese phonology. The TST comprised three sets, including (1) V and CV monosyllabic words and non-words, (2) V and CV trisyllabic non-words, and (3) trisyllabic words and non-words (the full stimulus list is attached as a supplementary file). Only the pure vowels (V) in the second set were analyzed in this study. The pure vowels in the second set included 36 items, which were derived from three V structures (i.e., [a], [ɔ], [ɛ]) and three Cantonese tones (i.e., T1, T2, T4). Vowel [ɛ] rather than [i] was selected as the high front vowel because [i] is also a high frequency vowel as [a] and [ɔ] in Cantonese (J. S. Li et al., 2023). If all high frequency vowels are included, it may raise the risk of overusing particular phonemes in the participants. Each item included all three tones, forming six different tone sequences: (1) T1T2T4, (2) T1T4T2, (3) T2T1T4, (4) T2T4T1, (5) T4T1T2, and (6) T4T2T1. All data collection sessions were conducted in sound-proof booth with the use of Audio-technica (model: AT2035) or Shure (model: SM48) microphones and Steinberg (Model: UR22mkII) or M-AUDIO (Model: M-Track Plus II) interfaces. The average time of completing the TST was about 45 minutes. Breaks were given to the participants upon request during the tasks.

2.3. Data analysis

The audio recordings were imported into Praat (Boersma & Weenink, 2022) for analysis. Following the procedure suggested by Ma et al. (2006), each vowel was identified manually from a wide-band spectrogram and an amplitude waveform visually displayed in the software. Segmentation and annotations were performed by a trained research assistant and a trained undergraduate student who were studying speech and language sciences. All of them received a two-hour training from a trainer, another research assistant, who had extensive experience in the acoustic analysis of Cantonese speakers. Before the annotators started independent analyses, all of them were required to obtain more than 0.80 on the Intraclass Correlation Coefficient (ICC) with the trainer on the f_0 values of an adult post-stroke (not included in this study) performing the TST. The ICC values obtained for training were 0.83 (95% CI: 0.82, 0.83, $p < .001$) and 0.85 (95% CI: 0.85, 0.86, $p < .001$), which were good reliability (Koo & Li, 2016). Intra- and inter-rater reliabilities on annotation were calculated with around 10% of total re-annotations performed by the annotators using two-way random effects, ICC (2, 1). The measurement variable was the f_0 values extracted from the annotation using the ProsodyPro (Xu, 2013), a custom-written f_0 analysis script for Praat. Excellent inter- (0.91 [95% CI: 0.91, 0.91, $p < .001$]), and intra-rater reliabilities (0.99 [95% CI: 0.99, 0.99, $p < .001$] and 0.96 ([95% CI: 0.96, 0.96, $p < .001$]) were obtained.

2.3.1. Vowel

F1 and F2 values were extracted using FormantPro (Xu & Gao, 2018), a custom-written formant analysis script for Praat. Midpoint formant values (i.e., F1 and F2 values extracted at

50% of the vowel duration) as well as the pairwise Euclidean distance between the midpoint of the three vowels (i.e., [ɛ]-[a], [ɛ]-[ɔ] and [ɔ]-[a]) to index static vocalic quality and vowel distinctiveness. Euclidean distance was calculated as the mean distance between two vowels on a two-dimension coordinate system using F1 and F2 values as axes. Within-speaker variation was assessed by coefficients of variance (CoVs), that is, to divide the standard deviation by the mean value of multiple repetitions of a given vowel by a participant for each tone and vowel sequences (Rogan et al., 1977). These measures test the influence of tones on vowel production (H1). Offset and onset formant values (i.e., F1 and F2 values extracted at 80% and 20% of the vowel duration) were used to investigate the vowel-to-vowel anticipatory (H2) and carryover co-articulating effects (H3) respectively.

Linear mixed-effect (LME) models were established for the formant values and CoVs of each vowel (i.e., [ɛ], [ɔ] and [a]) as well as the EDs between vowel pairs (i.e., [a]-[ɛ], [a]-[ɔ] and [ɛ]-[ɔ]) to evaluate the influence of *participant group* (AOS vs. nAOS vs. HC), *tone* (High-level vs. Rising vs. Falling), the *preceding* or *following vowel* (None vs. [a] vs. [ɛ] vs. [ɔ]), and *position* in the TST sequence (Initial vs. Middle vs. Final), using *glmer* in the *lmerTest* package (Kuznetsova et al., 2017) in R (RStudio Team, 2020). The optimal fixed structure of each model was selected by stepwise comparisons from the simplest structure to the most complex. Since we were particularly interested in the interaction between fixed effects, Type III analyses with Satterthwaite methods were conducted (Kuznetsova et al., 2017) to test the significance of an effect with all the other effects in the model (Kuznetsova et al., 2017; Satterthwaite, 1946). Tukey post hoc test was used to assess the significance of differences between pairs of group means (Abdi & Williams, 2010) with an emphasis on the interaction between *participant group* and other fixed effects. The details of the LME models are presented as appendices.

2.3.2. Tone

As previously mentioned, time-normalized f_0 values were extracted with ProsodyPro (Xu, 2013) for graphic comparison by dividing the duration of each tone-bearing vowel into 20 equal intervals and extracting f_0 values at each 5%. Offset and onset f_0 values extracted at 80% and 20% of the vowel duration were used to measure anticipatory (H2) and carryover (H3) co-articulating effects respectively.

LME models were established to evaluate the influence of *participant group*, the *preceding* or *following tone* (High-level vs. Rising vs. Falling) on the f_0 values through the same processes as the LME models on vocalic parameters. Models were established for each individual tone (i.e., high-level tone, rising tone, and low-falling tone) separately, and were further tested with Satterthwaite methods. Tukey test was also used for post-hoc comparison with an emphasis on the interaction between *participant group* and other fixed effects. The details of the LME models are presented as appendices.

3. Results

3.1. Vocalic quality at middle point

This section addresses H1 by presenting the measures of vocalic quality at middle points in each group, which reflect the influence of tones on vowel production in sequence. Vowel quadrilaterals at middle points are shown in Figure 2.

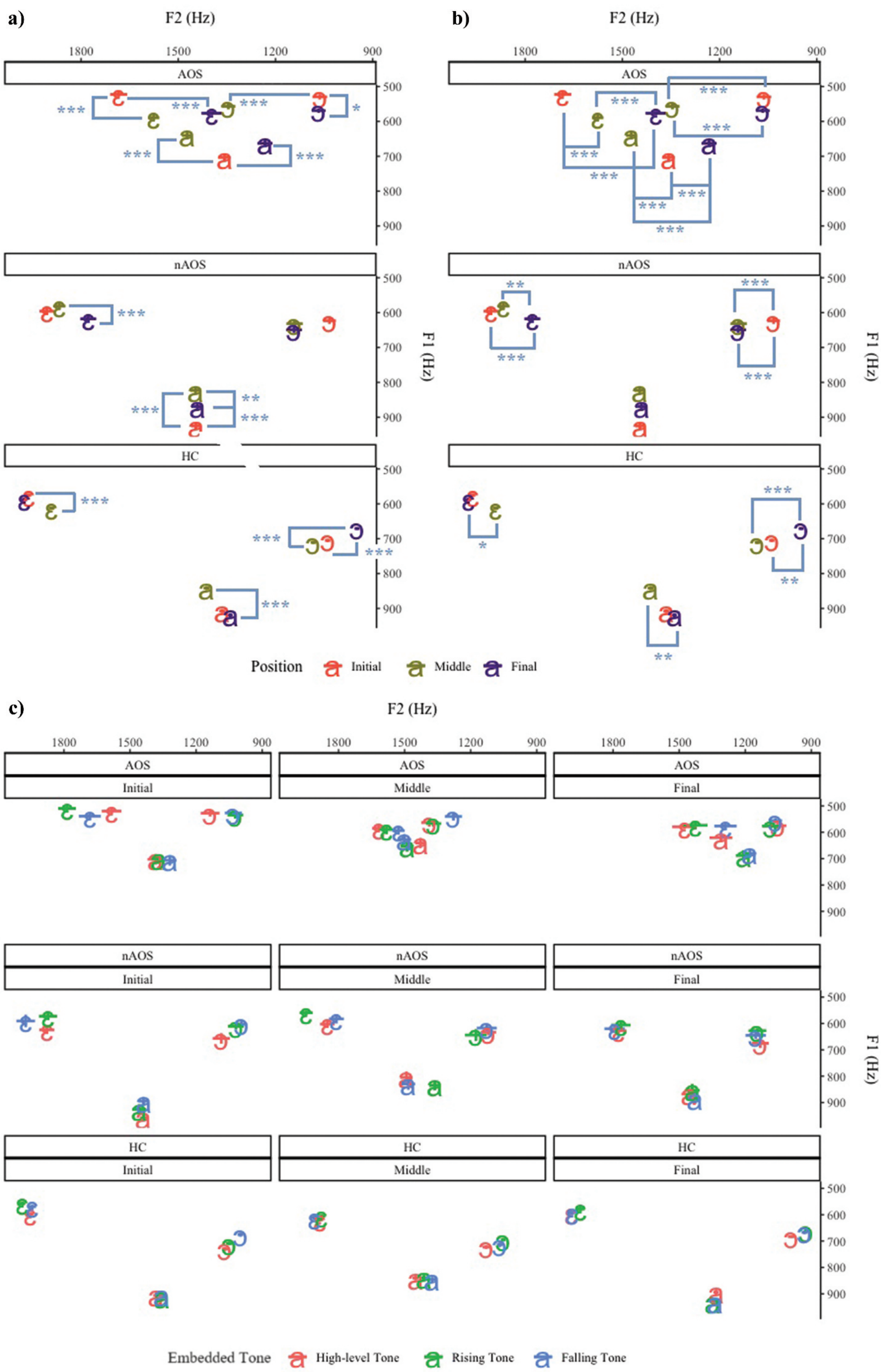

Figure 2. Middle point vowel quadrilaterals. (A) Separated by *participant group* and *position* with post-hoc comparison results of F1. (B) Separated by *participant group* and *position* with post-hoc comparison results of F2. C) separated by *participant group*, *tone* and *position*. [note. Statistically significant differences between specific comparisons are indicated in (A) and (B) by asterisk, in which * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$, while significant differences in C) were described in the text due the overlapping between vowels in the figure; AOS = AOS group; nAOS = nAOS group; HC = healthy control group].

3.1.1. Formant values

Evaluation of the LME models on F1 showed that the inclusion of the two-way interaction between *participant group*, and *position* significantly contributed to the models for all three vowels [ɛ] ($\chi^2(4) = 18.03$, $p < .05$), [ɔ] ($\chi^2(4) = 14.40$, $p < .001$), and [a] ($\chi^2(4) = 10.70$, $p < .001$), whereas the three-way interaction between *participant group*, *tone* and *position* also significantly contributed to the models for the two relatively back vowels [ɔ] ($\chi^2(8) = 2.27$, $p < .05$) and [a] ($\chi^2(8) = 2.53$, $p < .01$) (for full models see Appendix A).

Post-hoc comparison of the interactions found no significant differences between the three groups in vowels at the same position while carrying the same tones, as illustrated in Figure 2(a). However, within each group, some significant differences between positions were observed, except for [ɔ] produced by the nAOS group. As predicted in H1, the AOS group found it challenging to maintain as distinctive vowel height at non-initial positions as at initial positions – significant differences between *initial* and the other two positions were found for all three vowels in the AOS, with the relatively high [ɛ] and [ɔ] being the highest and the low [a] the lowest at the initial position; such pattern and significance were not systematically found in the other two control groups. Furthermore, the significant raising effects of the high-level tone on F1 values predicted in H1 were only found in the [a] at the final position in the AOS group (*p*s < .05). By contrast, some significant lowering effects of the high-level tone were observed in the [ɔ] at the initial position produced by the nAOS (*p*s < .05) and the HC groups (high-level vs. low-falling, $p < .05$) (Figure 2(c)). No statistically significant differences were found in the rest of the comparisons.

In terms of F2, the inclusion of the two-way interaction between *participant group*, and *position* significantly contributed to the LME models for all three vowels [ɛ] ($\chi^2(4) = 23.56$, $p < .001$), [ɔ] ($\chi^2(4) = 29.99$, $p < .001$) and [a] ($\chi^2(4) = 17.96$, $p < .001$), whereas the three-way interaction between *participant group*, *tone* and *position* significantly contributed to the models for [ɛ] ($\chi^2(8) = 3.68$, $p < .001$) and [a] ($\chi^2(8) = 3.94$, $p < .001$) (for full models see Appendix A). Similar to F1, there were more significant pairwise differences between positions in the AOS group than in the two control groups, indicating more variance in vowel frontness triggered by position and hence less stability in the vocalic quality produced by the AOS group, supporting H1 (Figure 2(b)). Such instability was also reflected in the fact that significant influence of tones on F2 was only seen in the AOS group, though the direction of the influence of tone at different positions was irregular – at initial position, [ɛ] carrying the high-level tone were significantly more back than [ɛ] carrying the rising tone ($p < .05$); by contrast, at final position, [ɛ] and [a] carrying the high-level tone were significantly more front than [ɛ] ($p < .05$) and [a] ($p < .01$) carrying the low-falling tone (Figure 2(c)). No statistically significant differences were found in the rest of the comparisons.

3.1.2. Euclidean distance (ED)

Evaluation of the LME models on EDs showed that the inclusion of the three-way interactions significantly contributed to the models for all three pairs, [ɛ]-[ɔ] ($\chi^2(4) = 3.42$, $p < .001$), [ɛ]-[a] ($\chi^2(4) = 2.61$, $p < .01$), and [ɔ]-[a] ($\chi^2(4) = 4.70$, $p < .001$) (Refer to Appendix B).

According to post-hoc comparisons, the EDs at the middle position produced by the AOS group, except the [ɛ]-[ɔ] carrying the low-falling tone, were all significantly smaller than those produced by the two control groups regardless of tones and positions (*p*s < .01). The [ɛ]-[a] carrying high-level tones produced at the final position by the AOS group was also significantly smaller than the other [ɛ]-[a] produced by the control groups (*ps* < .01). No statistically significant differences were found in the rest of the comparisons. The significant smaller EDs in the AOS group indicate less distinctive vowels produced by this group, supporting H1.

Within-group comparisons also found larger impacts of both positions and tones and hence more variance in vowel production by participants with AOS (Figure 2(c)). In terms of [ɛ]-[ɔ], in the AOS group, the ED carrying the high-level tone (vowels carrying high-level tone are red in colour in Figure 2) was significantly smaller than the ED carrying the rising tone (vowels carrying rising tone are green in colour in Figure 2) at the initial position. In the nAOS group, when carrying a low-falling tone (vowels carrying falling tone are blue in colour in Figure 2), the ED at the initial position was significantly larger than that in the middle ($p < .001$) and final ($p < .05$) positions. In terms of [ɛ]-[a], in the AOS group, the ED at the middle position carrying the falling tone was significantly smaller than all others except the ED at the final position carrying the high-level tone (vs. Initial: *ps* < .001; vs. Others: *ps* < .05). In the HC group, the EDs at the middle positions were significantly smaller than the EDs at the final positions carrying rising and falling tones (*ps* < .05). In terms of [ɔ]-[a], in the AOS group, the ED at the initial position carrying the high-level tones was significantly larger than the ED at the middle and final positions carrying the same tone (*ps* < .05), and the ED at the initial position carrying falling tone was also significantly larger than the EDs at the middle position carrying the rising and falling tones (*ps* < .001). No statistically significant differences were found in the rest of the comparisons.

3.1.3. Coefficients of variance (CoVs)

With regard to within-speaker variation, mean CoVs of F1 and F2 values in the three groups are presented in Table 2. Evaluation of the LME models on the CoV of F1 found significant influence of *participant group* ($\chi^2(2) = 7{,}87$, $p < .01$), *tone* ($\chi^2(2) = 9.83$, $p < .001$), the interaction between these two effects ($\chi^2(4) = 2.73$, $p < .05$) on [ɛ] as well as significant influence of *participant group* ($\chi^2(2) = 4.74$, $p < .05$), and *position* ($\chi^2(2) = 4.07$, $p < .05$) on [ɔ], but no significant influence of the fixed factors or their interaction was found in the F1 of the low vowel [a] (Appendix C). Post-hoc comparisons found that the AOS group

Table 2. Mean coefficients of variance (CoVs) of F1 and F2 values in three groups.

		CoV of F1 ± SD			
Group	Vowel	High-level tone	Rising tone	Low-falling tone	CoV of F2 ± SD
AOS	[ɛ]	7.12 ± 4.86	8.71 ± 4.64	9.52 ± 5.47	11.3 ± 9.33
	[ɔ]	9.56 ± 5.45			12.3 ± 9.70
	[a]	8.25 ± 5.64			9.02 ± 7.18
nAOS	[ɛ]	8.77 ± 10.4	8.73 ± 5.78	13.8 ± 12.5	8.94 ± 7.19
	[ɔ]	8.64 ± 8.00			10.9 ± 9.83
	[a]	9.25 ± 8.77			8.21 ± 7.84
HC	[ɛ]	4.32 ± 2.29	5.28 ± 2.90	5.57 ± 3.17	4.09 ± 3.48
	[ɔ]	4.86 ± 3.18			6.86 ± 4.28
	[a]	4.83 ± 3.48			5.1 ± 3.02

Notes: CoV = Coefficients of variance; AOS = AOS group; nAOS = nAOS group; HC = Healthy control group; SD = standard deviation. CoVs index within-speaker variations. The values are reported according to the results of LME models. The mean CoVs of all the vowels except the F1 of [ɛ] were calculated over groups and vowels, which were further separated by the tone carried by the vowels.

produced the largest CoV of F1 of [ɔ], significantly larger than the HC group ($p < .05$). However, with regard to [ɛ], the nAOS group produced the largest CoV of F1 when the tone was the low-falling tone, significantly larger than the [ɛ] bearing high-level and rising tones ($ps < .01$) as well as all the [ɛ] produced by HC ($ps < .01$); in fact, the CoV of F1 of the falling-tone bearing [ɛ] produced by the nAOS group was even larger than the [ɛ] bearing high-level tone produced by the AOS group, ($p < .05$). No statistically significant differences were found in the rest of the comparisons. The CoV of F2 was also significantly influenced by *participant group* in [ɛ] ($\chi^2(2) = 5.86$, $p < .05$) and [ɔ] ($\chi^2(2) = 5.21$, $p < .05$), but not [a] (Appendix C), with the AOS group producing the largest CoV, slightly larger than the nAOS group and significantly larger than the HC group (Table 2). As H1 predicted, the AOS group showed larger within-speaker variation than HC. Surprisingly, however, it was the nAOS rather than AOS showed a larger variation in the height dimension (i.e., CoVs of F1) than the frontness (CoVs of F2).

3.2. Vowel-to-vowel co-articulation

3.2.1. Anticipatory co-articulation between vowels

This section address H2 by presenting formant values at offsets under the influence of the following vowels. Vowel quadrilaterals at offsets are shown in Figure 3.

3.2.1.1. Vowel height. With regard to vowel height, evaluation of the LME models on the offset F1 showed that the inclusion of the two-way interaction between *participant group* and *following vowel* significantly contributed to the models for all three vowels [ɛ] ($\chi^2(4) = 11.10$, $p < .001$), [ɔ] ($\chi^2(4) = 23.98$, $p < .001$), and [a] ($\chi^2(4) = 14.67$, $p < .001$), but the three-way interaction between *participant group*, *following vowel* and *tone* only significantly contributed to the models for the low [a] ($\chi^2(8) = 1.95$, $p < .05$) (Refer to Appendix D).

Post-hoc comparisons were illustrated in Figures 3(a,b). In Figure 3(a), in [ɛ], significant differences between the following vowels were only seen in the nAOS group. However, different from the expected anticipatory co-articulating pattern, that is, vowels followed by lower vowels would have lower height, it was the [ɛ] followed by the low vowel [a] (blue [ɛ]) being the highest. In [ɔ], significant influence of the following vowels was seen in both nAOS and HC, but again, the HC group demonstrated the expected co-articulating pattern (i.e., [ɔ] followed by low [a] being the lowest) but not the nAOS group. In Figure 3(b), the final [a] (purple [a]) was the lowest in the two control groups, but significant differences induced by the following vowels were only found when in the falling tone condition in the nAOS but in both falling and rising tone conditions in the HC group. From Figure 3(a,b), it also could be seen that no statistical significance was found in the AOS group, indicating a highly limited influence of following vowels on the F1 values of the preceding vowels, and hence the lack of anticipatory co-articulation as predicted in H2.

3.2.1.2. Vowel frontness. Similarly, with regard to vowel frontness, the inclusion of the two-way interaction between *participant group* and *following vowel* significantly contributed to the LME models on the offset F2 for all three vowels [ɛ] ($\chi^2(4) = 34.10$, $p < .001$), [ɔ] ($\chi^2(4) = 20.99$, $p < .001$) and [a] ($\chi^2(4) = 31.66$, $p < .001$), but the three-way interaction only significantly contributed to the models for the low [a] ($\chi^2(8) = 6.52$, $p < .05$) (Refer to Appendix D).

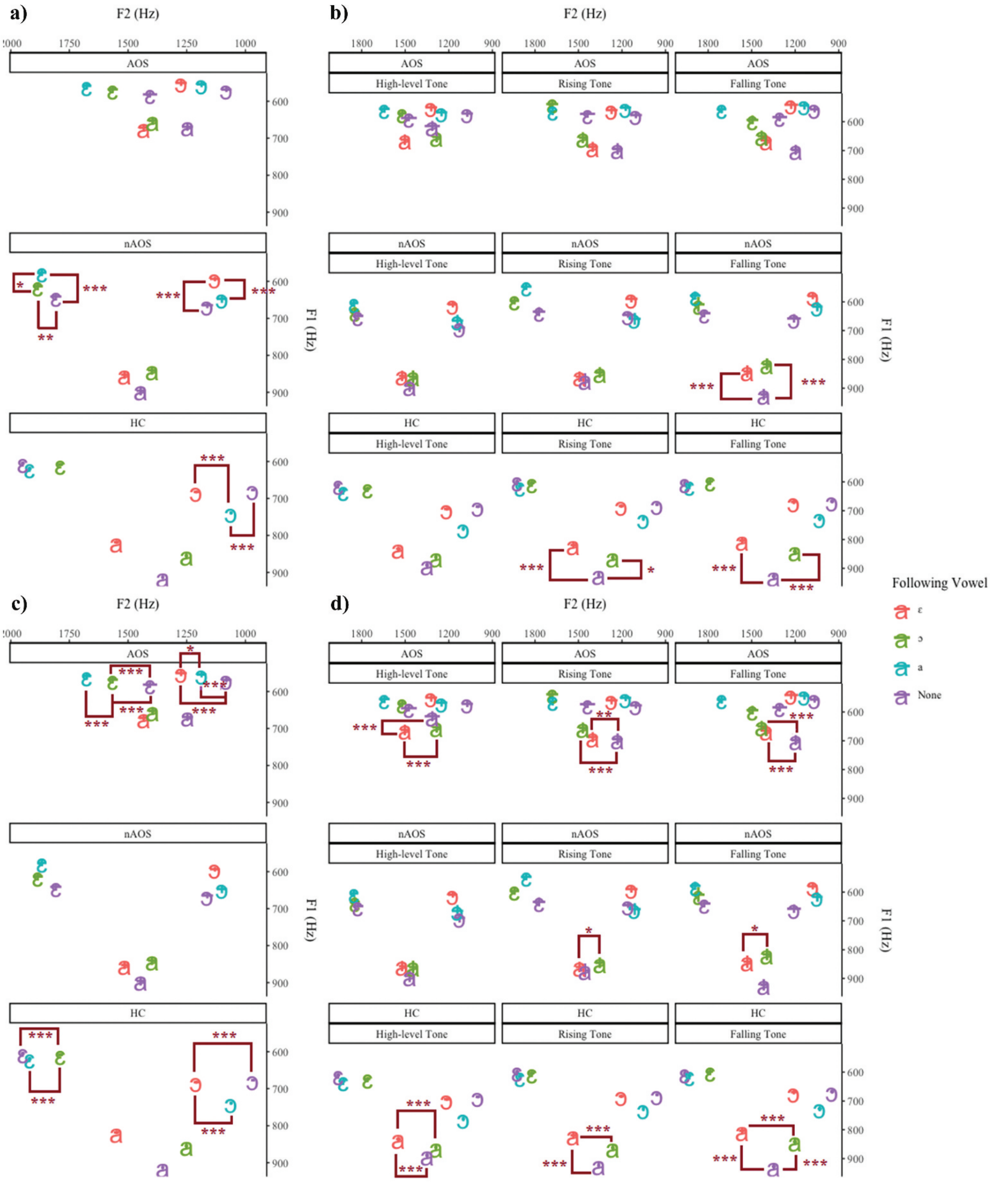

Figure 3. Offset vowel quadrilaterals (A) separated by *participant group* and *following vowel* with post-hoc comparison results of F1. (B) Separated by *participant group*, *tone*, and *following vowel* with post-hoc comparison results of F1. C) separated by *participant group* and *following vowel* with post-hoc comparison results of F2. D) separated by *participant group*, *tone* and *following vowel* with post-hoc comparison results of F2. [note. Statistically significant differences between specific comparisons are indicated by asterisk, in which * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$; AOS = AOS group; nAOS = nAOS group; HC = healthy control group.].

Post-hoc comparisons were illustrated in Figures 3(c,d). In Figure 3(c), in [ɛ] and [ɔ], the final vowel (purple) was significantly more back than those followed by other vowels in the AOS group, while in the HC group, as expected, it was the back vowel [ɔ] (green) that significantly moved back [ɛ] and [ɛ] that significantly moved front [ɔ] (red); although no statistical significance was seen in the nAOS group, this group did not show a regular co-articulating pattern either. In Figure 3(d), in terms of [a], tones only changed the degree rather than direction of anticipatory co-articulation in the two control groups, i.e., with [a] followed by the vowel [ɔ] (green [a]) being the most back and [a] followed by the vowel [ɛ] (red [a]) being the most front; by contrast, in the AOS group, the direction of anticipatory co-articulation also differed with tones, as only when carrying high-level tone, speakers with AOS produced the same co-articulating patterns as the nAOS and HC groups. Overall, unlike in the HC group, not much similarization was seen between the preceding and the following vowels the AOS group in the frontness dimension, supporting H2.

3.2.2. Carryover co-articulation between vowels

This section address H3 by presenting formant values at onsets under the influence of the preceding vowels. Vowel quadrilaterals at onsets are shown in Figure 4.

3.2.2.1. Vowel height. With regard to vowel height, evaluation of the LME models on the onset F1 showed that the inclusion of the two-way interaction between *participant group* and *preceding vowel* significantly contributed to the models for all three vowels [ɛ] ($\chi^2(4) = 22.73$, $p < .001$), [ɔ] ($\chi^2(4) = 11.66$, $p < .001$), and [a] ($\chi^2(4) = 17.07$, $p < .001$), but the three-way interaction between *participant group*, *following vowel* and *embedded tone* only significantly contributed to the models for the back [ɔ] ($\chi^2(8) = 3.47$, $p < .001$) (Refer to Appendix E).

Post-hoc comparisons were illustrated in Figures 4(a,b). In Figure 4(a), in [ɛ] and [a], the AOS group showed regular carryover co-articulation, as [ɛ] followed by [a] (blue [ɛ]) was the lowest while [a] followed by [ɛ] (red [a]) was the highest, and differences were of statistical significance; regular patterns were found in the HC group, but post-hoc comparisons only found statistical significance in [a] but not [ɛ]. By contrast, in Figure 4(b), the direction of co-articulating patterns of [ɔ] differed between tones in the AOS and nAOS groups, while only in the HC group was the carryover co-articulation of [ɔ] regular and statistically significant. Carry-over co-articulation between vowels in the height dimension in the AOS group was more regular than anticipatory co-articulation, but not as regular as that in the HC group, partly supporting H3.

3.2.2.2. Vowel frontness. With regard to frontness (Figure 4(c,d), evaluation of the LME models on the onset F2 showed that the inclusion of the two-way interaction between *participant group* and *preceding vowel* significantly contributed to the models for all three vowels [ɛ] ($\chi^2(4) = 4.95$, $p < .001$), [ɔ] ($\chi^2(4) = 13.10$, $p < .001$) and [a] ($\chi^2(4) = 9.83$, $p < .001$), and the three-way interaction between *participant group*, *following vowel* and *embedded tone* only significantly contributed to the models for the two relatively front vowel [ɛ] ($\chi^2(8) = 2.15$, $p < .05$) and [a] ($\chi^2(8) = 3.13$, $p < .01$) but not [ɔ] (Refer to Appendix E).

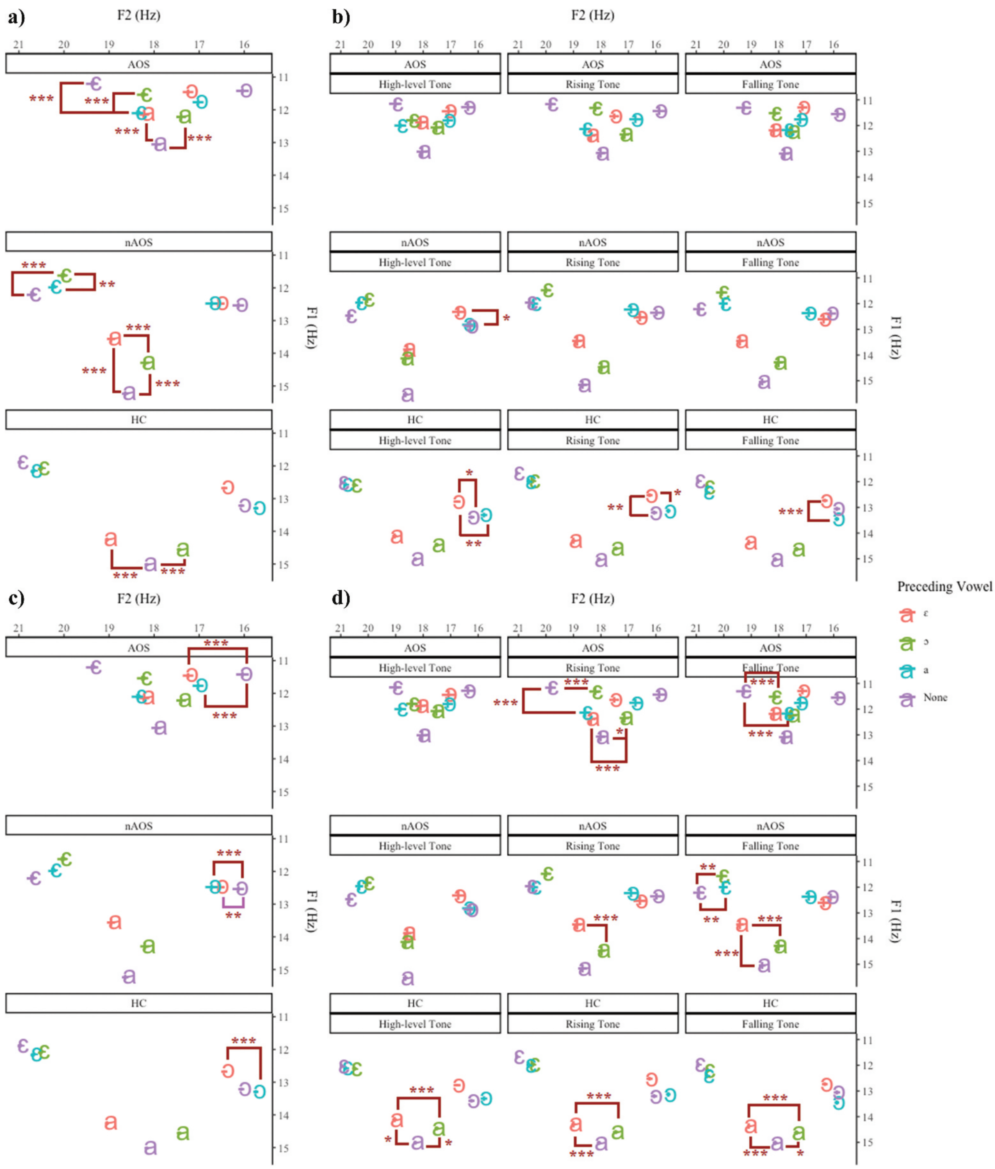

Figure 4. Onset vowel quadrilaterals (A) separated by *participant group* and *preceding vowel* with post-hoc comparison results of F1. (B) Separated by *participant group*, *tone* and *preceding vowel* with post-hoc comparison results of F1. C) separated by *participant group* and *preceding vowel* with post-hoc comparison results of F2. D) separated by *participant group*, *tone* and *preceding vowel* with post-hoc comparison results of F2. [note. Statistically significant differences between specific comparisons are indicated by asterisk, in which * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$; AOS = AOS group; nAOS = nAOS group; HC = healthy control group].

Regular carryover co-articulation was found in all three groups in terms of the F2 (Figure 4(c)); tones only influence the degree of co-articulation, and hence the statistical significance in [ɛ] and [a] (Figure 4(d)). In line with H3, carry-over co-articulation in the height dimension was regular in the AOS group, similar to those observed in the controls.

3.3. Tonal co-articulation

Tone production in the AOS group was far less accurate than the other two groups, as the AOS participants tended to level the first two tones and make the final tone fall regardless of the intrinsic shapes of the tones (Figure 5, top row). Figure 6 showed the offset and onset f_0 values of each tone in different tonal contexts.

3.3.1. Anticipatory Co-articulation between tones

This section addresses H2 by presenting f_0 offset under the influence of the following tones. Evaluation of the LME models showed that the inclusion of the two-way interaction between *tone* and *following tone* significantly contributed to the models for the offset of all three tones: the high-level tone ($\chi^2(4) = 24.19$, $p < .001$), the rising tone ($\chi^2(4) = 9.88$, $p < .01$), and the low-falling tone ($\chi^2(4) = 25.56$, $p < .001$) (Refer to Appendix F).

Post-hoc comparisons were illustrated in Figure 6. As predicted in H2, not much anticipatory co-articulation between tones was seen in the AOS group. In Figures 6(a-1), no anticipatory effect was seen in the offset of the high-level tones, and the AOS group (red lines) demonstrated a changing pattern that is opposite to the two control groups in the offsets of the rising and falling tones, mainly due to the final falling. In addition, in the AOS group, the offsets of the two contour tones at final were significantly lower than those followed by other tones (Figures 6(a-2)), and the offset of the falling tone was not

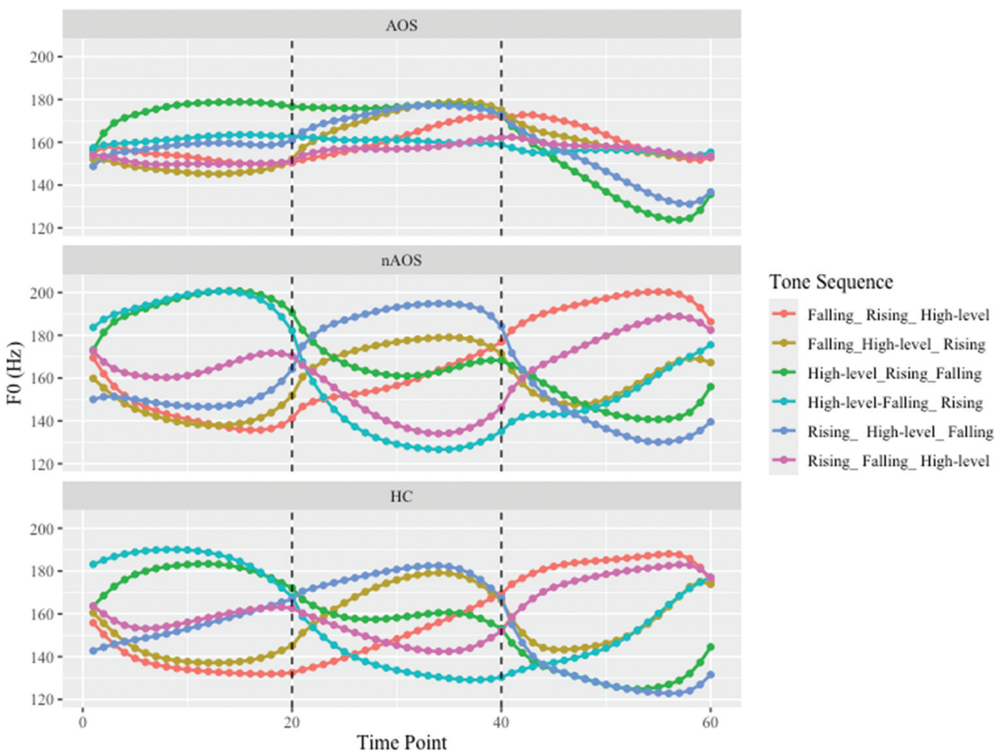

Figure 5. Time-normalized tonal contours produced by the three participant groups. [note. AOS = AOS group; nAOS = nAOS group; HC = healthy control group].

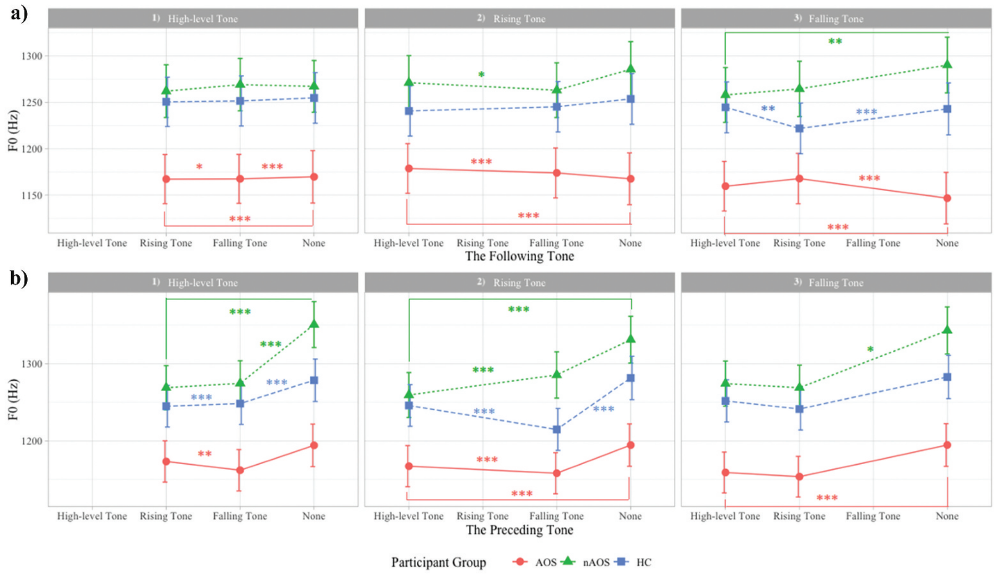

Figure 6. Tonal co-articulation. (A) Offset f_0 values with post-hoc comparison results. (B) Onset f_0 values with post-hoc comparison results. [Note. Statistically significant differences between specific comparisons are indicated by asterisk, in which * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$; AOS = AOS group; nAOS = nAOS group; HC = healthy control group].

raised by the following high-level tone but by the rising tone (Figures 6(a-3)). By contrast, the rising tone in the nAOS (green lines) and HC (blue lines) groups reached the highest target at the final position (Figures 6(a-2)), while the offset of the low-falling tone was raised by the following high-level tone but lowered by the mid-onset of the following rising tone (Figures 6(a-3)). No significant pre-low raising effect was observed in tones preceding the low-falling tone in the present study.

3.3.2. *Carryover co-articulation between tones*

This section addresses H3 by presenting f_0 onset under the influence of the preceding tones. Similarly, the inclusion of the three-way interaction between *embedded tone* and *following tone* significantly contributed to the models for all three tones, the high-level tone ($\chi^2(4) = 11.08$, $p < .001$), the rising tone ($\chi^2(4) = 14.28$, $p < .001$), and the falling tone ($\chi^2(4) = 6.40$, $p < .001$) (Refer to Appendix F). Unlike the anticipatory co-articulation, the carryover co-articulation in the AOS group (red lines) was of similar regularity to that in the HC group (blue lines), as predicted in H3. The onsets of the initial tones ("none" in Figures 6(b1-3)) were the highest due to the lack of down-step in all three groups, while the preceding low-falling tone lowered the onset of the following high-level and rising tone in the AOS and HC groups but not the nAOS group (green lines) ("Falling Tone" in Figures 6(b-2) and 6(b–3)).

4. Discussion

This study investigated the co-articulation between vowel and tone, vowel and vowel, and tone and tone in Cantonese speakers with and without AOS. Section 4.1 summarises the main findings of the study, and further discussion in comparison with the existing literature is included in 4.2 and 4.3.

4.1. Main findings

Analyses of the midpoint vowel quality partially support H1. The AOS group produced significantly different vowels compared to the nAOS and HC groups, including less distinctive vowels at non-initial positions, especially when vowels carried high-level and rising tones, and less stable vowel quadrilateral under the influence of the position in the TST sequence and the tones they carried. However, not much raising effect of high tone on F1 was found except in the final [a] produced by the AOS group, different from H1. Also, unlike suggested in the previous studies (e.g., Jacks et al., 2010), the height dimension indexed by F1 in the AOS group did not show larger deviations or within-speaker variation than the frontness dimension indexed by F2; in fact, some significant differences were seen in the frontness dimension, e.g., the non-initial front [ɛ] produced by the AOS group was more backwards than those produced by the HC group. The Cantonese speaking participants with AOS in the present study may have also simplified F1 performance when simplifying tone (f_0) realization, leading to less variation in F1, the formant closer to f_0. Since Jacks et al. (2010) investigated non-tonal language speakers, the large deviations in F1 found in their study may not be applicable to participants who speak tone languages.

Analyses of anticipatory co-articulation between vowels and between tones support H2. The AOS group showed disrupted anticipatory co-articulation between vowels and between tones, compared to the nAOS and HC groups. In particular, the HC group showed regular vocalic and tonal anticipatory co-articulation, while the AOS and nAOS showed irregular anticipatory co-articulation between vowels. Differences between the AOS and nAOS groups in vowel-to-vowel co-articulation are observed in different dimensions, namely, the significant differences found in the AOS group were mainly observed in F2, the frontness dimension, while those found in the nAOS were mainly seen in F1, the height dimension. Regarding tonal anticipatory co-articulation, the AOS group demonstrated a final lowering pattern, which was not observed in both the nAOS and HC groups.

Results also support our H3, as assimilation to the previous vowels and tones was observed in the onset formants of vowels as well as the f_0 of tones in all three groups of participants, indicating a regular carryover co-articulation in the AOS and the control groups.

4.2. Vowel and tone productions

One of the clinical features that we found in speakers with AOS post-stroke is “increased errors with increased length or complexity” (Duffy et al., 2023, p. 484). In the present study, we found vowel deviations and tone simplifications in the AOS groups, especially at the non-initial positions in a sequencing context. Such findings confirmed the difficulty in the simultaneous production of segments and tones in Cantonese speakers with AOS post-stroke. The results are consistent with W. Chen et al. (2022), who demonstrated that Mandarin speakers with AOS have both vowel and tone disruptions. Our findings extended that Cantonese speakers with AOS may also have difficulty in producing vowels with tones in the sequencing context, which requires continuous and simultaneous variation of vowel and tone. Furthermore, the present results show that the contour tones expanded the vowel quadrilaterals at the initial positions but brought distortions and compression at the middle and final positions. When speakers with AOS need to continuously and simultaneously manage multiple parameters (such as tone and vowel), they may prioritize one parameter over another as the sequence progress.

We propose that the simplification of tones in the AOS group may be due to their undershooting tonal targets. It is also possible to explain the simplification from a motor perspective. The AOS group may have deficits of simultaneous motor planning of segments and tones that might not be found in the control groups. According to he GODIVA model (Bohland et al., 2010), such deficits indicate that speakers with AOS have a deviated phonological content buffer but an intact sequential structure buffer in the planning loop. Furthermore, Maas et al. (2008) identified two types of motor preparation for speech articulation: the preprogramming stage of processing and the online serial ordering process, arguing that speakers with AOS have deficits in the former, a stage sensitive to the complexity of the unit, but not the latter. The present finding also echoes with Maas et al. (2008) in that the simultaneous realisation of tone and vowels leverages the complexity of the units in Cantonese; hence, the speakers with AOS are forced to realise the syllables with deviation or even substitution so as to maintain the number of syllables. In other words, to reduce the complexity, the Cantonese speakers with AOS post-stroke replaced all the tones in the final position by a falling tone, as falling is the physiologically easiest contour (Sundberg, 1979). It is worth mentioning, however, the simplification of tones does not improve vowel performance in the AOS group. Although they tended to level the initial and middle tones while making the final tone a falling tone, the intended tone shapes still disrupted their vowel formants, especially the F2 of [a], at the final positions. The larger distortion seen in the non-initial position also suggests a relatively ease initiation of movement.

4.3. Co-articulation

The present study also found problematic anticipatory co-articulation in both vowels and tones in the AOS group. Such findings are in line with the existing literature that reported delayed and reduced anticipatory co-articulation between segments among speakers with AOS who speak non-tonal languages like English and German (e.g., Southwood et al., 1997; Ziegler & von Cramon, 1986). The finding confirms that anticipatory co-articulation between suprasegmentals (i.e., tones in this case) is as problematic in the population with speech motor planning impairment as that between segments. It further adds to the evidence that tone is part of a word's motor programme (i.e., the planning of tones occurs at the motoric level), as suggested in the DIVA (Guenther, 2016) and GODIVA (Bohland et al., 2010) models. However, to what extent tones, elements that are traditionally regarded as suprasegmentals in the phonological literature (Halle & Vergnaud, 1982; Yip, 2002b), are like segments in the motor programme of words calls for more cross-linguistic evidence. Unlike in Cantonese, where almost all morphemes are monosyllabic and carry a single tone, in Shanghai Wu dialects and Bantu languages, the contour of each tone operates at the word level, while words can be of multiple syllables (Y. Chen & Gussenhoven, 2015; Goldsmith, 1988).

Furthermore, the more disrupted anticipatory co-articulation than the carryover co-articulation found in the AOS group included in this study also added to the evidence that anticipatory co-articulation reflects speech planning more than the carry-over effects (Recasens, 1984; Whalen, 1990). In addition, disrupted anticipatory co-articulation was also seen in the height dimension of vowels in the nAOS, though their anticipatory co-articulation in the frontness dimension and between tones was comparable to the HC groups. Such distortion may arise from the dysarthria two participants in the nAOS group had. Nevertheless, the AOS group did not show larger distortion in the height dimension indexed by F1 than the frontness, probably because their realization of tones was no less problematic as previously explained. In this way, their F1, the formant closer to f_0 was not as much impacted. The discrepancy between the AOS and nAOS also suggested that the problematic anticipatory co-articulation in Cantonese (and tone-vowel interaction in general) was still mainly due to AOS rather than the co-existing dysarthria.

The TST tones produced by the AOS and nAOS groups were also less accurate than those produced by the HC controls as judged by healthy native Cantonese speakers, even though the acoustic analyses found relatively little difference in the f_0 values between the nAOS and HC groups (Wong, Wong, Chen, et al., 2024). The irregular co-articulation found in the present study may help to explain the reduced tone accuracy of the two post-stroke groups found by Wong, Wong, Chen, et al. (2024). The perceptual judgment of tone is not influenced by a single acoustic correlate (e.g., f_0 and duration), but by the interaction between them (J. T. Gandour & Harshman, 1978). The problematic isochronism and therefore unnaturalness caused by the irregular co-articulation may lead native listeners to rate the tones produced by the speakers with and without AOS post-stroke as incorrect. The co-articulation in post-stroke speakers of Cantonese calls for further research as it may help to differentiate Cantonese speakers with AOS and aphasia from Cantonese speakers without AOS but with aphasia and from healthy speakers.

4.4. Clinical implications

This study demonstrated that Cantonese speakers with AOS post-stroke can be differentiated from those without AOS post-stroke and HC by observing vowel disruptions and tone simplifications. We focused on acoustic analyses in this study, but in another study (Wong, Wong, Chen, et al., 2024), we also reported the percentage of tones correct (PTC). PTC is calculated based on the listeners' perceptual judgment of tone correctness. Both previous and this study showed that pitch variational measures in a sequencing context may be useful in identifying AOS in tonal language speakers. In addition, our results demonstrated that Cantonese speakers with AOS may have difficulty in producing vowels and tones simultaneously, which may not be applicable to non-tonal language speakers. It indicates that the clinical manifestations of AOS in different language speakers may be different. For clinicians working with tonal language speakers with AOS, it is important to address both segmental and tonal aspects in the assessment and treatment process. For clinicians working with tonal language speakers with AOS, it is important to address both segmental and tonal aspects in the assessment and treatment process.

4.5. Limitations and future investigations

The present study is a preliminary study investigating the co-articulation between tones and vowels, between vowels and between tones in post-stroke speakers with AOS and with and without aphasia. Only five participants were included in each group. A larger sample size is recommended in future studies so as to generate more representative results for generalization. In addition, three participants in the AOS group also had dysarthria. Although no abnormality in pitch gliding was reported from the participants, the difficulty in laryngeal functions including reduced sustained phonation, limited loudness variation, and atypical voice quality, though mild, may contribute to the abnormal performance of the AOS group found in the present study. It is worth mentioning that the present findings can still be attributed to the presence of AOS, since two nAOS participants were also diagnosed of dysarthria of similar severity. It is also recognised that people with AOS are likely to have other speech disorders like dysarthria, but in future studies with a larger sample size, the clinical background of the participants could be better controlled. Furthermore, the present study only analyses sequences with isolated vowels without onset or coda consonants. Since both V and CV stimuli are used in TST-based assessment, analyses of CV and mixed sequences could be done in the future to examine the interaction between consonants, vowels, and tones. Lastly, the present study only conducted acoustic analyses. Whether the observed differences influence perception calls for further research, and the results may help to deepen our understandings of the communication profile of the population with AOS.

5. Conclusion

In summary, this preliminary analysis demonstrates that the AOS group had difficulty producing vowels with tones (i.e., both segments and tones), while the nAOS group produced vowels and tones similar to the HC group but with larger variations. Such difficulties did not only influence the production accuracy of certain segmental or suprasegmental entities, but also the coarticulation between them. A possible explanation is that the speakers with AOS post-stroke have difficulty with simultaneous management of segments and tones. This explanation is consistent with another study that found that Cantonese speakers with AOS post-stroke have difficulty in simultaneous management of syllable structure and tone in a tone sequencing context (Wong, Wong, Chen, et al., 2024).

Acknowledgements

The authors expressed their gratitude to all the participants and their families for their participations. The authors would also like to thank the student helpers and research assistants who performed the data analyses and reliability measures, including Ms. Twinkie Cheung, Ms. Helen Chung, Ms. Carly Ng, Mr. Ventus Siu, Ms. Dahlia Tang, and Ms. Cherry Yu. Appreciation was also given to the SLP students, Mr. Zenas Chan and Ms. Joyce Lin, for their participation in data collection and to two SLPs, Mr. Wilbert Chan and Mr. Jason Kan, for their viewing of the assessment videos and making diagnoses.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Dean's Reserve for Research, Scholarly and Other Endeavours of the Faculty of Humanities, the Hong Kong Polytechnic University Project ID: [P0036018]. This study was also supported by the Research Grants Council Early Career Scheme of the Hong Kong SAR Government (Project Number 25609819), granted to M. N. Wong.

ORCID

Yixin Zhang http://orcid.org/0000-0002-0485-2230
Eddy C.H. Wong http://orcid.org/0000-0003-1827-6876
Min Ney Wong http://orcid.org/0000-0003-0585-3566

References

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. *Encyclopedia of Research Design*, *3*(1), 1–5.

Ainsworth, W. A. (1975). Intrinsic and extrinsic factors in vowel judgments. In G. Fant (Ed.), *Auditory analysis and perception of speech*. Elsevier.

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, & 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, & Hearing Research*, *55*(6), 1822–1835. https://doi.org/10.1044/1092-4388(2012/11-0257)

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

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*, *22*(7), 1504–1529. https://doi.org/10.1162/jocn.2009.21306

Boutsen, F. R., & Christman, S. S. (2002). Prosody in apraxia of speech. *Seminars in Speech and Language*, *23*(4), 245–256. https://doi.org/10.1055/s-2002-35799

Brunelle, M. (2009). Tone perception in northern and southern Vietnamese. *Journal of Phonetics*, *37*(1), 79–96. https://doi.org/10.1016/j.wocn.2008.09.003

Chang, Y. C., & Hsieh, F. F. (2012). Tonal coarticulation in Malaysian hokkien: A typological anomaly? *The Linguistic Review*, *29*(1), 37–73. https://doi.org/10.1515/tlr-2012-0003

Chen, S., Wiltshire, C., & Li, B. (2018). An updated typology of tonal coarticulation properties. *Taiwan Journal of Linguistics*, *16*(2), 79–114. https://doi.org/10.6519/TJL.2018.16(2).3

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), 1–24. https://doi.org/10.1080/02699206.2022.2081611

Chen, Y., & Gussenhoven, C. (2015). Shanghai Chinese. *Journal of the International Phonetic Association*, *45*(3), 321–337. https://doi.org/10.1017/S0025100315000043

Choi, J. D., & Keating, P. (1991). Vowel-to-vowel coarticulation in three Slavic languages. *UCLA Working Papers in Phonetics*, *78*(78), 78–86.

Duffy, J. R. (2020). The disorders and their diagnoses: Apraxia of speech. In J. R. Duffy (Ed.), *4th*, motor speech disorders: Substrates, differential diagnosis, and management (pp. 269–193). Elsevier Health Science.

Duffy, J. R., Martin, P. R., Clark, H. M., Utianski, R. L., Strand, E. A., Whitwell, J. L., & Josephs, K. A. (2023). The apraxia of speech rating scale: Reliability, validity, and utility. *American Journal of Speech-Language Pathology*, *32*(2), 469–491. https://doi.org/10.1044/2022_AJSLP-22-00148

Enderby, P., Palmer, R. (2008). *Frenchay dysarthria assessment-second edition (FDA-2)*. Speechmark Publishing.

Erickson, D., Iwata, R., Endo, M., & Fujino, A. (2004, March). Effect of tone height on jaw and tongue articulation in mandarin Chinese. *Paper presented at the International Symposium on Tonal Aspects of Languages: With emphasis on tone languages*, *Beijing, China*.

Gandour, J., Potisuk, S., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., & Boongird, P. (1996). Tonal coarticulation in Thai after unilateral brain damage. *Brain and Language*, *52*(3), 505–535. https://doi.org/10.1006/brln.1996.0027

Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech*, *21*(1), 1–33. https://doi.org/10.1177/002383097802100101

Goldsmith, J. (1988). Prosodic trends in the Bantu languages. In N. Smith, H. van der Hulst (Eds.), *Autosegmental studies in pitch accent* (pp. 81–93). Foris.

Gu, W., & Lee, T. (2009). Effects of tone and emphatic focus on F0 contours of cantonese speech: A comparison with standard Chinese. *Chinese Journal of Phonetics*, *2*, 133–147.

Guenther, F. H. (2016). *Neural control of speech*. The MIT Press.

Haley, K. (2004). Vowel duration as a cue to postvocalic stop voicing in aphasia and apraxia of speech. *Aphasiology*, *18*(5–7), 443–456. https://doi.org/10.1080/02687030444000200

Haley, K. L., Cunningham, K. T., Jacks, A., Richardson, J. D., Harmon, T., & Turkeltaub, P. E. (2021). Repeated word production is inconsistent in both aphasia and apraxia of speech. *Aphasiology*, *35*(4), 518–538. https://doi.org/10.1080/02687038.2020.1727837

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

Haley, K. L., Jacks, A., Richardson, J. D., & Wambaugh, J. L. (2017). Perceptually salient sound distortions and apraxia of speech: A performance continuum. *American Journal of Speech-Language Pathology*, *26*(2S), 631–640. https://doi.org/10.1044/2017_AJSLP-16-0103

Haley, K. L., Ohde, R. N., & Wertz, R. T. (2000). Single word intelligibility in aphasia and apraxia of speech: A phonetic error analysis. *Aphasiology*, *14*(2), 179–201. https://doi.org/10.1080/026870300401540

Haley, K. L., Ohde, R. N., & Wertz, R. T. (2001). Vowel quality in aphasia and apraxia of speech: Phonetic transcription and formant analyses. *Aphasiology*, *15*(12), 1107–1123. https://doi.org/10.1080/02687040143000519

Halle, M., & Vergnaud, J. R. (1982). On the framework of autosegmental phonology. In H. van Der Hulst & N. Smith (Eds.), *The structure of phonological representations* (Vol. 2). Foris.

Han, M. S., & Kim, K. O. (1974). Phonetic variation of Vietnamese tones in disyllabic utterances. *Journal of Phonetics*, *2*(3), 223–232. https://doi.org/10.1016/S0095-4470(19)31272-0

Hombert, J. M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, *55*(1), 37–58. https://doi.org/10.2307/412518

Hoole, P., & Hu, F. (2004). *Tone-vowel interaction in standard Chinese*. Paper presented at the International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing, China.

Jacks, A., Mathes, K. A., & Marquardt, T. P. (2010). Vowel acoustics in adults with apraxia of speech. *Journal of Speech, Language, & Hearing Research*, *53*(1), 61–74. https://doi.org/10.1044/1092-4388(2009/08-0017)

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research*, *26*(2), 231–249. https://doi.org/10.1044/jshr.2602.231

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lee, A., Prom-On, S., & Xu, Y. (2021). Pre-low raising in cantonese and Thai: Effects of speech rate and vowel quantity. *Journal of the Acoustical Society of America*, *149*(1), 179–190. https://doi.org/10.1121/10.0002976

Lee, K. Y. S. (2012). *Hong Kong cantonese tone identification Test. Department of otorhinolaryngology, head and neck surgery*. The Chinese University of Hong Kong.

Li, B., Guan, Y., & Chen, S. (2020). Carryover effects on tones in Hong Kong Cantonese. In *10th International Conference on Speech Prosody*, Tokyo, Japan (pp. 489–493). https://doi.org/10.21437/850SpeechProsody.2020-100

Li, J. S., Badrulhisham, H., & Alderete, J. (2023). *Lexical and sub-lexical frequency effects in cantonese*. https://www.sfu.ca/~alderete/pubs/liEtal2020_cantfreq.pdf

Lindblom, B. E., & Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America*, *42*(4), 830–843. https://doi.org/10.1121/1.1910655

Ma, J. K.-Y., Ciocca, V., & Whitehill, T. L. (2006). Effect of intonation on cantonese lexical tones. *Journal of the Acoustical Society of America*, *120*(6), 3978–3987. https://doi.org/10.1121/1.2363927

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language*, *106*(2), 107–118. https://doi.org/10.1016/j.bandl.2008.03.004

Magen, H. S. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *25*(2), 187–205. https://doi.org/10.1006/jpho.1996.0041

Manuel, S. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 179–198). Cambridge University Press.

Manuel, S. Y., & Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research*, *77*(78), 69–78.

McNeil, M. R., Ballard, K. J., Duffy, J. R., & Wambaugh, J. L. (2017). Apraxia of speech theory, assessment, differential diagnosis, and treatment: Past, present, and future. In P. van Lieshout, B. Maassen, & H. Terband (Eds.), *Speech motor control in normal and disordered speech: Future developments in theory and methodology* (pp. 195–221). ASHA Press.

Mok, P. K. P., & Hawkins, S. (2004, December). Effects of phonemic vs allophonic density and stress on vowel-to-vowel coarticulation in cantonese and Beijing Mandarin. In *2004 International symposium on Chinese spoken language processing*, Hong Kong, China (pp. 33–36). IEEE.

Mok, P. P. (2012). Effects of consonant cluster syllabification on vowel-to-vowel coarticulation in English. *Speech Communication*, *54*(8), 946–956. https://doi.org/10.1016/j.specom.2012.04.001

Mok, P. P. (2013). Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech*, *56*(2), 191–209. https://doi.org/10.1177/0023830912443948

Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1991). Perceptual characteristics of vowel and prosody production in apraxic, aphasic, and dysarthric speakers. *Journal of Speech, Language, & Hearing Research*, *34*(1), 67–80. https://doi.org/10.1044/jshr.3401.67

Ohala, J. J. (1973). The physiology of tone. L. M. Hyman (Eds), *Southern california occasional papers in linguistics* (Vol. 1, pp. 1–14).

Ostry, D. J., Gribble, P. L., & Gracco, V. L. (1996). Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned? *Journal of Neuroscience*, *16*(4), 1570–1579. https://doi.org/10.1523/JNEUROSCI.16-04-01570.1996

Parmenter, C. E., Tevino, S. N., & Bevans, C. A. (1933). The influence of a change in pitch on the articulation of a vowel. *Language*, *9*(1), 72–81. https://doi.org/10.2307/409515

Potisuk, S., Gandour, J., & Harper, M. P. (1997). Contextual variations in trisyllabic sequences of Thai tones. *Phonetica*, *54*(1), 22–42. https://doi.org/10.1159/000262208

Qin, Z., & Zhang, J. (2022). The use of tonal coarticulation cues in cantonese spoken word recognition. *JASA Express Letters*, *2*(3). https://doi.org/10.1121/10.0009728

Recasens, D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. *Journal of the Acoustical Society of America*, *76*(6), 1624–1635. https://doi.org/10.1121/1.391609

Rogan, J. C., Keselman, H. J., & Breen, L. J. (1977). Assumption violations and rates of type I error for the Tukey multiple comparison test: A review and empirical investigation via a coefficient of variance variation. *The Journal of Experimental Education*, *46*(1), 20–25. https://doi.org/10.1080/00220973.1977.11011605

RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC. https://www.rstudio.com/

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110–114. https://doi.org/10.2307/3002019

Shen, X. S. (1990). Tonal coarticulation in mandarin. *Journal of Phonetics*, *18*(2), 281–295. https://doi.org/10.1016/S0095-4470(19)30394-8

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

Southwood, M. H., Dagenais, P. A., Sutphin, S. M., & Garcia, J. M. (1997). Coarticulation in apraxia of speech: A perceptual, acoustic, and electropalatographic study. *Clinical Linguistics & Phonetics*, *11*(3), 179–203. https://doi.org/10.3109/02699209708985190

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

Sundberg, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics*, *7*(2), 71–79. https://doi.org/10.1016/S0095-4470(19)31040-X

Torng, P. C. (2000). *Supralaryngeal articulator movements and laryngeal control in mandarin Chinese tonal production* [Doctoral dissertation]. University of Illinois at Urbana-Champaign.

Torng, P. C., Alfonso, P., & van Lieshout, P. (2000). Lexical tone specification on vocal tract shape in mandarin Chinese vowel production. In *Proceedings of the 5th Speech Production Seminar*, Seeon, Germany (pp. 73–76).

Utianski, R. L., Duffy, J. R., Clark, H. M., Strand, E. A., Boland, S. M., Machulda, M. M., Whitwell, J. L., & Josephs, K. A. (2018). Clinical progression in four cases of primary progressive apraxia of speech. *American Journal of Speech-Language Pathology*, *27*(4), 1303–1318. https://doi.org/10.1044/2018_AJSLP-17-0227

van der Merwe, A. (2021). New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology*, *35*(4), 397–423. https://doi.org/10.1080/02687038.2020.1765306

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

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, *18*(1), 3–35. https://doi.org/10.1016/S0095-4470(19)30356-0

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

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

Wong, E. C. H., Wong, M. N., & Velleman, S. L. (2024). Acoustic analysis of tone productions in a sequencing context among Cantonese-speaking preschool children with childhood apraxia of speech. *Journal of Speech, Language, & Hearing Research*, *67*(6), 1682–1711. https://doi.org/10.1044/2024_JSLHR-23-00383

Xu, Y. (1997). Contextual tonal variations in mandarin. *Journal of Phonetics*, *25*(1), 61–83. https://doi.org/10.1006/jpho.1996.0034

Xu, Y. (2013). *ProsodyPro-A tool for large-scale systematic prosody analysis*. Laboratoire Parole et Langage. https://discovery.ucl.ac.uk/id/eprint/1406070

Xu, Y., & Gao, H. (2018). FormantPro as a Tool for Speech Analysis and Segmentation / FormantPro como uma ferramenta para a análise e segmentação da fala. *Revista de Estudos da Linguagem*, *26*(4), 1435–1454. https://doi.org/10.17851/2237-2083.26.4.1435-1454

Yip, M. (2002a). Introduction: How is tone produced? In M. Yip (Ed.), *Tone* (pp. 5–10). Cambridge University Press.

Yip, M. (2002b). *Tone*. Cambridge University Press.

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

Zee, E. (1980). Tone and vowel quality. *Journal of Phonetics*, *8*(3), 247–258. https://doi.org/10.1016/S0095-4470(19)31474-3

Ziegler, W., & von Cramon, D. (1986). Disturbed coarticulation in apraxia of speech: Acoustic evidence. *Brain and Language*, *29*(1), 34–47. https://doi.org/10.1016/0093-934x(86)90032-5