

Articulatory–Acoustical Relationship in Cantonese Vowels*

Wai-Sum Lee

City University of Hong Kong

Language and Linguistics
17(4) 477–500
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1606822X16637058
lin.sagepub.com



This paper investigates the articulatory–acoustical relationship in the production of the Cantonese long point vowels [i: u: a:] in CV syllables and shortened equivalents [i u a] in checked CVC syllables, through an analysis of the tongue positions and corresponding formant frequencies for the vowels. The articulatory and acoustical data from four Cantonese speakers, two male and two female, show that: (i) the shortened high vowels [i] and [u], relative to the long [i:] and [u:], undergo formant undershoot in the acoustical vowel space and [u] also undergoes target reduction in the articulatory vowel space, and (ii) the effect of shortened vowel duration is not observed in the low [a]. The data on the articulatory–acoustical relationship show that: (i) the levels of sensitivity of F_1 and F_2 to the variations in tongue displacement are low in each of the vowels [i: u: a:] and [i u a], suggesting that the Cantonese point vowels are quantal, and (ii) the sensitivity of the formant frequencies is lower to the variation in tongue constriction in the x -dimension than in the y -dimension in the long [i: u:], but not the long [a:] and shortened [i u a]. This suggests that: (a) the low vowels are less quantal than the high vowels and (b) the shortened vowels are less quantal than the long equivalents. The results are discussed in connection with the theory of vowel target reduction, the quantal theory, and the articulatory–acoustical relationship in the English point vowels.

Key words: articulatory–acoustical relationship, Cantonese vowels, quantal theory, theory of vowel target reduction

1. Introduction and background

The theory of vowel reduction proposed in Lindblom (1963) predicts that formants undergo undershoot in the acoustical vowel space when vowel duration decreases, as ‘the speech organs fail, as a result of the physiological limitations, to reach the positions that they assume when the vowel is pronounced under ideal steady-state conditions’ (Lindblom 1963:1779). Since the advent of the theory, there have been a number of studies that explore the effect of duration on vowel production in different languages. Behne et al. (1996) investigated the effect of vowel duration on the spectral characteristics of the phonologically distinct long and short vowels, [i: O: A:] and [i O A] (sic), in Norwegian. Speech data from six male and six female speakers showed that the difference in

* An earlier version of this paper was presented at the 14th International Symposium on Chinese Languages and Linguistics, Taipei, 4–6 June 2014. This research was supported by a GRF grant (9041707) from the Research Grants Council of the Government of Hong Kong SAR. I am grateful to the two anonymous reviewers for their constructive comments on this paper. Special thanks go to Mark Tiede of the Haskins Laboratories for providing me with the scripts for using EMMA.

vowel duration between the long [i: O: A:] (approximately 130–150 ms) and the short [i O A] (approximately 80–90 ms) does not have a significant effect on the spectral characteristics of the short vowels, with only a minimum difference in F_1 or F_2 between any pair of the long and short vowels. Johnson & Martin (2001) investigated the effect of vowel duration on the formant frequencies for two sets of phonologically distinct long [i: o: a:] and short [i o a] vowels in Creek, a Muskogean Amerindian language. Speech data from four male and four female speakers showed that the mean duration of the long vowels [i: o: a:] (225 ms) is almost twice as long as the short equivalents [i o a] (124 ms). Relative to the long ones, the short vowels undergo formant undershoot in the acoustical vowel space in that F_1 is larger for [i o] than [i: o:] and smaller for [a] than [a:], while F_2 is smaller for [i] than [i:] and larger for [o] than [o:]. It was further reported that formant undershoot in the short vowels is also found in many quantity languages, such as Serbo-Croatian, Czech, Hungarian, Cairo Arabic, and Scottish Gaelic, in which the vowels contrast in length, and there are cross-language variations in degree of vowel reduction in the short vowels. Mooshammer & Geng (2008), a study of formant and articulatory target undershoot in the German vowels, examined the effect of a change in vowel length on the long tense vowels [i: y: e: ε: ø: a: o: u:] and the short lax vowels [ɪ ʏ ε œ a ɔ ʊ] in unstressed position. Speech data from seven speakers, five male and two female, showed that in the unstressed position, the degree of vowel shortening is pronounced in the tense vowels, but minimal in the lax ones. The shortening of the tense vowels results in a noticeable reduction in the articulatory vowel space, with a forward shift of the back vowels and an upward shift of the low vowels, but no noticeable change in the front vowels. As for the lax vowels, the reduction is small. Spectrally, the acoustical vowel space is reduced for both the tense and lax vowels in the unstressed positions, as a result of the raising of the low vowels (a decrease in F_1) and centralization of both the front (a decrease in F_2) and back vowels (an increase in F_2). Thus, in the unstressed positions, both the front tense and lax vowels undergo formant undershoot in the acoustical vowel space, but not vowel target reduction in the articulatory vowel space. Mooshammer & Geng (2008) attributed the difference to the possibility that there are other articulatory modifications aside from the change in tongue position.

Tabain (2003a, 2003b) and Tabain & Perrier (2005, 2007) investigated the French point vowels [i u a] that occur in the pre-boundary, that is the final position, of different prosodic units, such as sentence, intonational phrase, accentual phrase, word, and syllable. Results based on the speech of three speakers showed that the effect of the lengthened vowel duration in pre-boundary position on articulation as well as F_1 and F_2 is observed during [a] (Tabain 2003a, 2003b), but not [i] (Tabain & Perrier 2005). Tabain & Perrier (2005, 2007) reported that during [i] and [u] the effect of duration on formants is similar for different speakers, but the effect on the tongue and lip positions varies across the speakers. That between speakers the different articulatory actions result in similar acoustic outputs suggests that the articulatory–acoustical relationship in the vowels is not linear.

Perkell & Nelson (1985), Perkell & Cohen (1989), and Beckman et al. (1995) investigated the variability of tongue displacement in the x - and y -dimensions in the English vowels and the relationships between formant frequencies and variations in tongue constriction location in the x -dimension and tongue constriction degree in the y -dimension. Similar findings were reported, in that the variability, which varies across the vowel types, is greater in tongue constriction location in the x -dimension than in tongue constriction degree in the y -dimension. The articulatory and

spectral data on the English point vowels [i u ɑ] generally support the claims made by Stevens' (1972, 1989) quantal theory. This predicts that, in some regions of the vocal tract, small shifts in articulation result in large changes in acoustical properties while, in other regions, the acoustic parameter remains unperturbed when large articulatory shifts are made. Furthermore, within the acoustically stable regions (Stevens 1972, 1989), where tongue constriction along the vocal tract is narrow, formant frequencies are relatively less sensitive to variation in tongue constriction location in the *x*-dimension than variation in tongue constriction degree in the *y*-dimension. However, the correlation between formant frequencies and variation in tongue displacement was found to be significant only in Beckman et al. (1995), but not Perkell & Nelson (1985) and Perkell & Cohen (1989). Beckman et al. (1995) further reported that the articulatory–acoustical relationships in the English point vowels [i] and [u] were found to be quantal, in that the relationship between articulatory movements and acoustical consequences in vowel production is non-linear, exhibiting acoustic stability when large articulatory shifts are made (Stevens 1972, 1989). Such a relationship is also observed in [æ], but not [ɑ]. It was concluded that since the vowels vary across languages, 'the extent to which a speech sound is quantal seems to be a variable property that depends not just on its distinctive feature specification, but also on its precise specification for habits of production for a given language or dialect' (Beckman et al. 1995:489).

Hoole (1999b) examined the relationship between formant frequencies and variations in tongue constriction location and tongue constriction degree for the German vowels [i e y ø] using electromagnetic articulography. Results based on the articulatory data from six male speakers showed that the correlations between F_2 and variation in tongue constriction location are close to none. This indicates that the palatal vowels are acoustically insensitive to perturbation in tongue constriction location, but the correlations between F_2 and variation in tongue constriction degree are strong. The results confirm what was reported in Beckman et al. (1995): that formant frequencies are relatively less sensitive to variation in tongue constriction location in the *x*-dimension than variation in tongue constriction degree in the *y*-dimension. As for F_1 and variation in tongue constriction degree, the correlation between them is also strong for the German palatal vowels, but the slope of the regression line is flatter for the rounded [y ø] than for the unrounded [i e].

Previous studies have been concerned with vowel reduction and the articulatory–acoustical relationship in the vowels of languages other than Chinese. The present paper considers how a difference in vowel duration affects the articulation of Cantonese vowels and the formant frequencies. In Cantonese, the seven long vowels [i: y: ε: œ: a: ɔ: u:] occur in CV syllables, and in checked CVC syllables they are shortened, symbolized as [i y ε œ a ɔ u] (Zee 1999). This paper investigates how a change in vowel duration affects the articulatory–acoustical relationship in the two sets of Cantonese point vowels, [i: u: a:] and [i u a]. It determines: (i) whether the shortened Cantonese vowels [i u a] undergo articulatory vowel reduction and formant undershoot relative to the long equivalents [i: u: a:], as would be predicted by the theory of vowel reduction (Lindblom 1963); and (ii) whether the Cantonese point vowels are similar to the English vowels for which (a) the degree of variability is larger in tongue constriction location in the *x*-dimension than in tongue constriction degree in the *y*-dimension, and (b) the formant frequencies are relatively less sensitive to variation in tongue constriction location than variation in tongue constriction degree (Beckman et al. 1995).

2. Method

2.1 Test materials

Table 1 presents the six Cantonese test monosyllabic words that contain one of the long point vowels [i: u: a:] or the shortened equivalents [i u a]. The vowels [i: u: a:] occur in open CV syllables and [i u a] in checked CVC syllables, where C is a zero-consonant, glottal fricative, or bilabial stop. The use of the test syllable [p^hut-] is compelled by the phonotactic constraint in Cantonese, where [u] in checked CVC syllables is obligatorily followed by [t].

Table 1: Test monosyllabic words used in the study

CV syllables		CVC syllables	
i:	[i:] 'to heal' (醫)	i	[hip-] 'to assist' (協)
u:	[u:] 'black' (烏)	u	[p ^h ut-] 'to splash' (潑)
a:	[ha:] 'shrimp' (蝦)	a	[hap-] 'to sip' (呷)

2.2 Speakers

Four native speakers of Cantonese, two male and two female, provided the speech and articulatory data. They were university students in their early twenties and had lived their whole life in Hong Kong. No speakers reported a history of speech or hearing problems. They were paid a small sum for participation in the experiment.

2.3 Data collection and analysis

Twenty repetitions of each test word were recorded by each speaker, making up a total of 480 test tokens (6 test words × 20 repetitions × 4 speakers) for analysis. The movements of the articulators and synchronized audio signals during the test words were digitally recorded using the Electromagnetic Mid-sagittal Articulography (EMMA) AG500 (Carstens Medizinelektronik, Göttingen, Germany). For background information on the technique and details of operation and setup of EMMA, see Perkell et al. (1992), Gracco (1995), and Hoole (1996).

To collect the data on tongue contours and tongue positions during the test vowels, as shown in Figure 1, three sensors (or receiver coils) were mounted on three different locations on the speaker's tongue: tongue-front, tongue-middle, and tongue-back, equidistantly along the midline of the tongue surface. Three additional sensors, one mounted on the nose bridge and the other two on the back of the ears served as fixed reference points for tracking head movements during recording. To obtain the mid-sagittal contour of the palate, a sensor was mounted on the tip of the thumb of the speaker, who was instructed to swipe the thumb forward, against and along the midline of the palate from approximately the point at which the hard palate ends to the back of the upper frontal incisors. For verification of the palate contour, dental impressions of the roof of the mouth were made for the speakers.

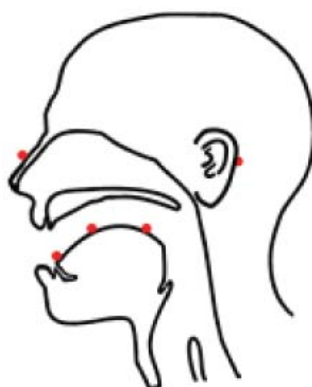


Figure 1: Positions of the sensors mounted on the tongue-front, tongue-middle, and tongue-back and the fixed reference sensors on the nose bridge and the back of each ear

2.4 Data analysis

The MATLAB scripts were used on the raw articulatory data for: (i) filtering the noise which was transmitted to the sensors during recording; and (ii) correcting the head movements. A data visualization and analysis tool, MVIEW, was employed to display audio signal, palate contour, spectrogram, and tongue contour.

The synchronized audio signals of the test words, which were digitally recorded through the built-in microphone of EMMA at a sampling rate of 16,000 Hz, were down-sampled to 10,000 Hz, producing an upper frequency cut-off of 5,000 Hz for analysis of vowel formants. Using the speech analysis software, Computerized Speech Lab (CSL4500; KayPENTAX), a pitch synchronized autocorrelation-based LPC (Linear Prediction Coding) spectral analysis was performed on the test vowels for the measurement of formant frequencies. Following Hoole (1999a) and Mooshammer & Geng (2008), the measurements were made at the vowel target or the temporal mid-point of the steady-state section of the formant trajectories for the vowels, which is considered to be ‘close to the point that would be exacted by means of a minimum articulatory velocity criterion’ (Hoole 1999a:1023). Note that all the test Cantonese vowels, long or shortened, in this study have a clearly identifiable steady-state section.

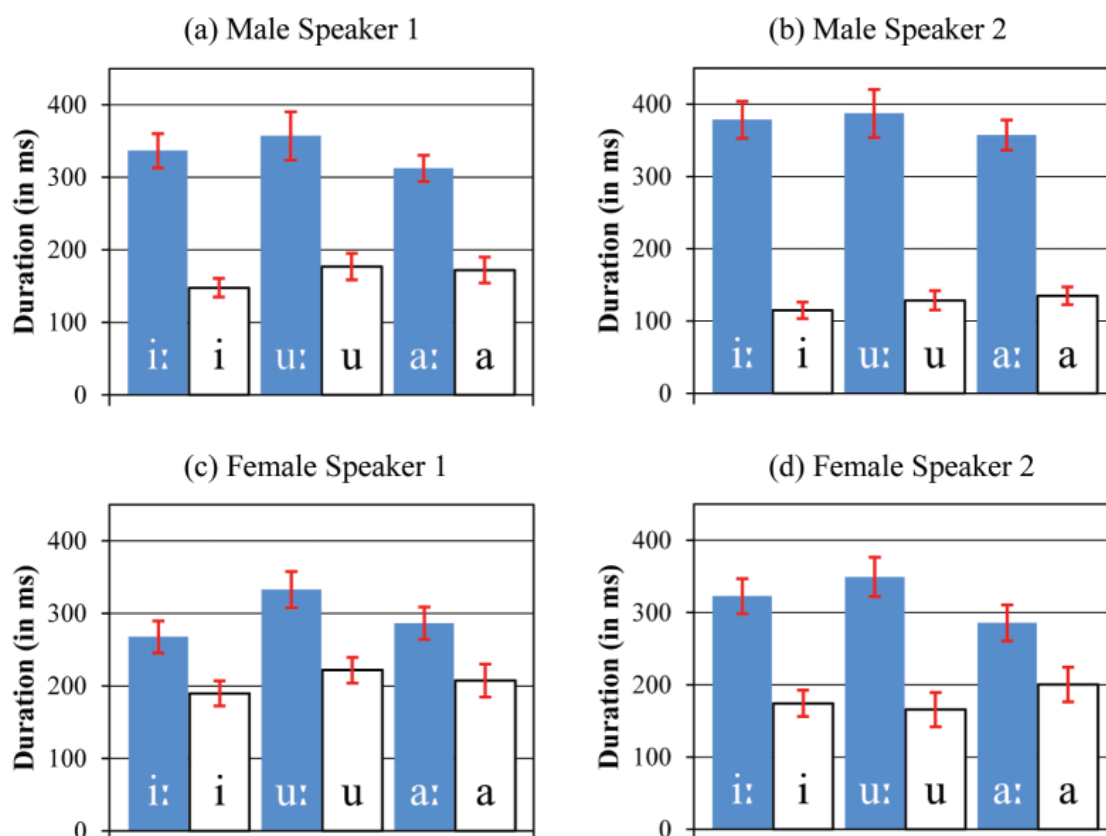
The time points selected for spectral analysis were also the time points at which the articulatory data on the tongue contours and tongue positions during the vowels were obtained. The duration (in ms) of each test vowel was measured directly from the speech waveforms. Since any test vowel in the present study was preceded and/or followed by a zero- or voiceless consonant, the vowel onset was set at the beginning of the first normal regular glottal pulse and the vowel offset at the end of the last well-defined glottal pulse.

3. Results

3.1 Duration

Figures 2a–2d display the bar graphs that show the durations (in ms) of the two sets of Cantonese vowels, long [i: u: a:] and shortened [i u a], for two male speakers (Figures 2a and 2b)

and two female speakers (Figures 2c and 2d). For each speaker, the duration of each vowel type is the average of 20 repetitions. Also shown in the figures are error bars representing one standard deviation of the mean duration of each vowel.



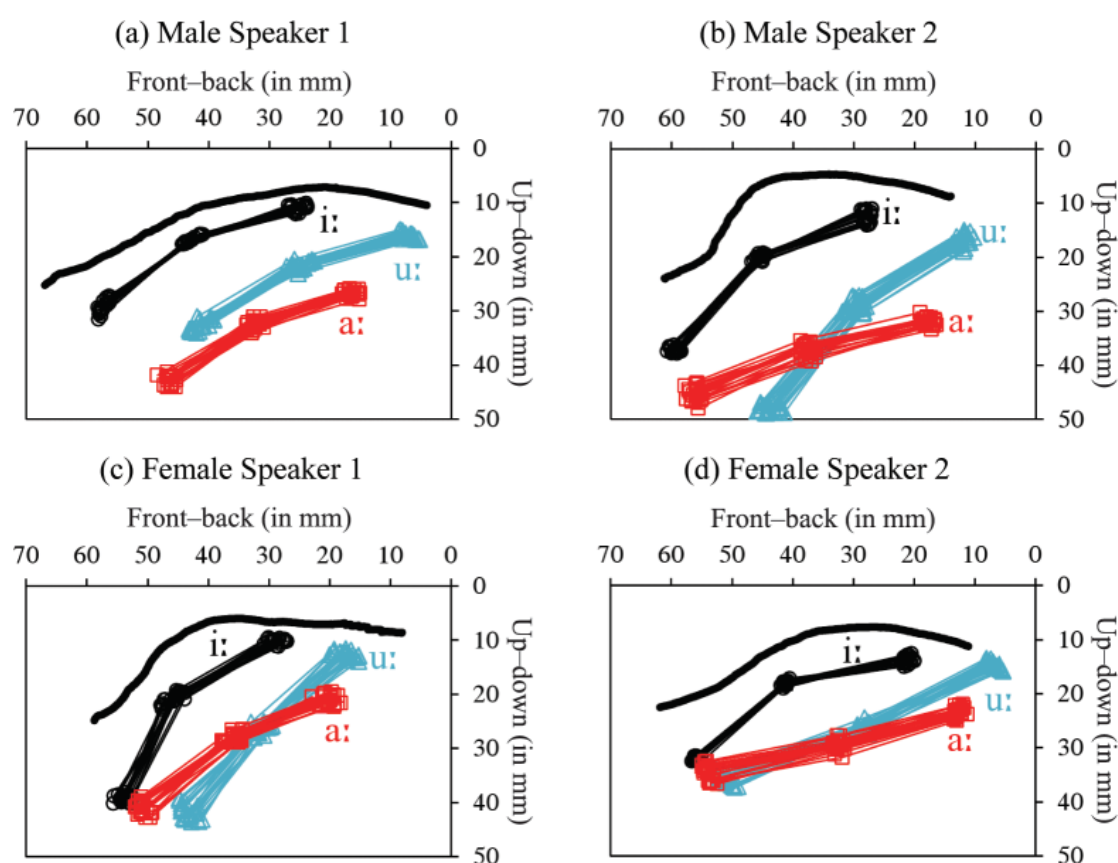
Figures 2a–2d: Mean duration (in ms; $n = 20$) and error bar representing one standard deviation of the mean duration of each one of the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

As can be seen in Figures 2a–2d, the durations of [i: u: a:] are markedly longer than those of [i u a] for all speakers. There are cross-speaker variations in duration between the two sets of vowels. The grand mean durations of the three vowels of each set show that the difference in duration between [i: u: a:] and [i u a] is larger for Male Speaker 2 (248.29 ms; Figure 2b) and smaller for Female Speaker 1 (89.41 ms; Figure 2c) than Male Speaker 1 (169.83 ms; Figure 2a) and Female Speaker 2 (138.99 ms; Figure 2d). The results of *t*-tests show that the difference between any pair of long and shortened vowels of the same vowel type is significant ($p < 0.0001$) for each speaker. Averaging across four speakers, the duration of the long vowels [i: u: a:] (331.16 ms) is 1.95 times longer than the shortened equivalents [i u a] (169.53 ms).

3.2 Articulatory characteristics

Figures 3a–3d show the superimposed mid-sagittal tongue contours and tongue positions for 20 repetitions of each one of the Cantonese long vowels [i: u: a:] for Male Speakers 1 and 2

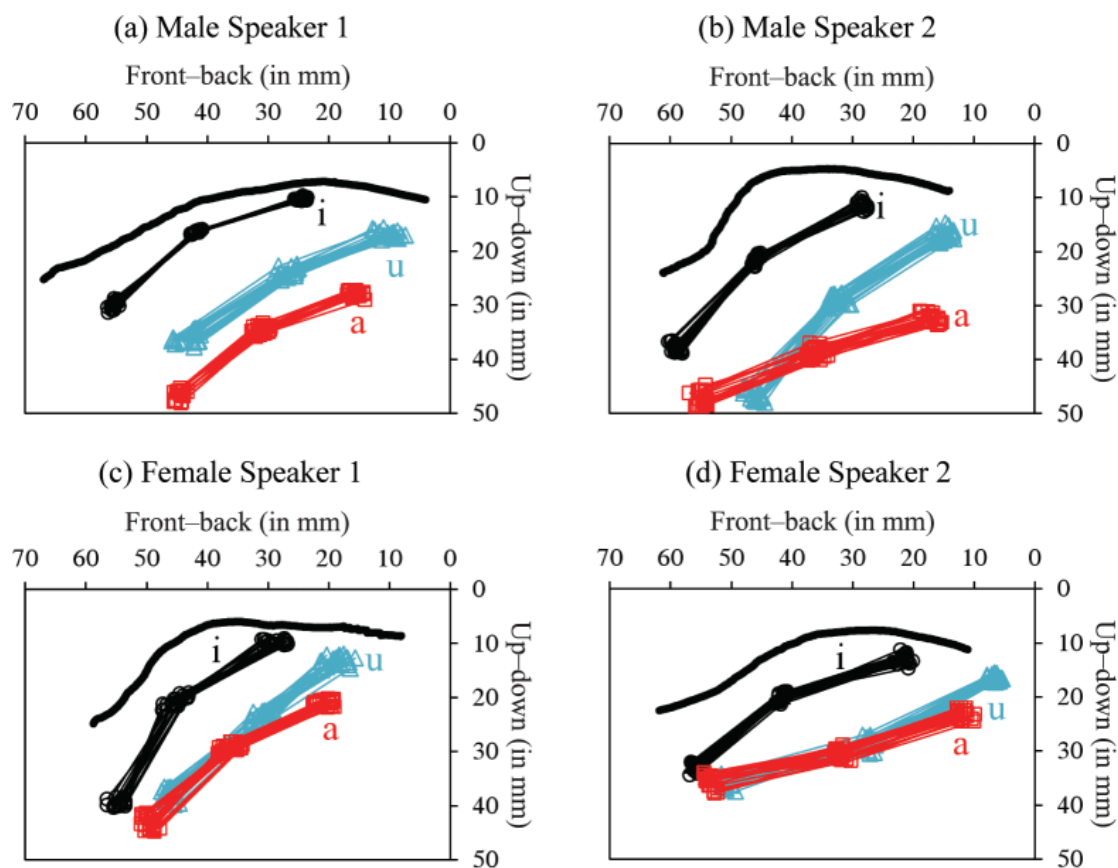
(Figures 3a and 3b) and Female Speakers 1 and 2 (Figures 3c and 3d). The tongue contours were obtained by connecting the three articulatory data points, that is tongue-front, tongue-middle, and tongue-back, with the front-back position on the *x*-axis against the up-down position on the *y*-axis. The figures are on the same scale (in mm) for all speakers, with the origin at the top right corner of the figures. An increase in value on the *x*- or *y*-axis denotes a forward shift or a lowering of the tongue position. The thick dark line in the figures is the tracing of the mid-sagittal palate contour serving as a reference, relative to which the tongue positions in both the front-back (*x*) and up-down (*y*) dimensions during the vowels are defined and compared.



Figures 3a–3d: Tongue contours and tongue positions for the Cantonese long vowels [i: u: a:] and the palate contours (thick dark line) for two male and two female speakers

As can be seen in Figures 3a–3d, the patterns of the tongue contour and tongue position in the mid-sagittal vocal tract for the three long vowels are similar across the four speakers. During the high front vowel [i:], the articulatory data point at the tongue-back position, that is the posterodorsum of the tongue, is close to the hard palate, or close to the soft palate during the high back vowel [u:]. As for the low vowel [a:], the tongue body is distant from the palate.

The superimposed mid-sagittal tongue contours and tongue positions for 20 repetitions of each one of the Cantonese shortened vowels [i u a] are shown in Figures 4a and 4b for Male Speakers 1 and 2 and Figures 4c and 4d for Female Speakers 1 and 2. Similar patterns in the tongue contour and tongue position are observed for the shortened vowels [i u a] across the four speakers.

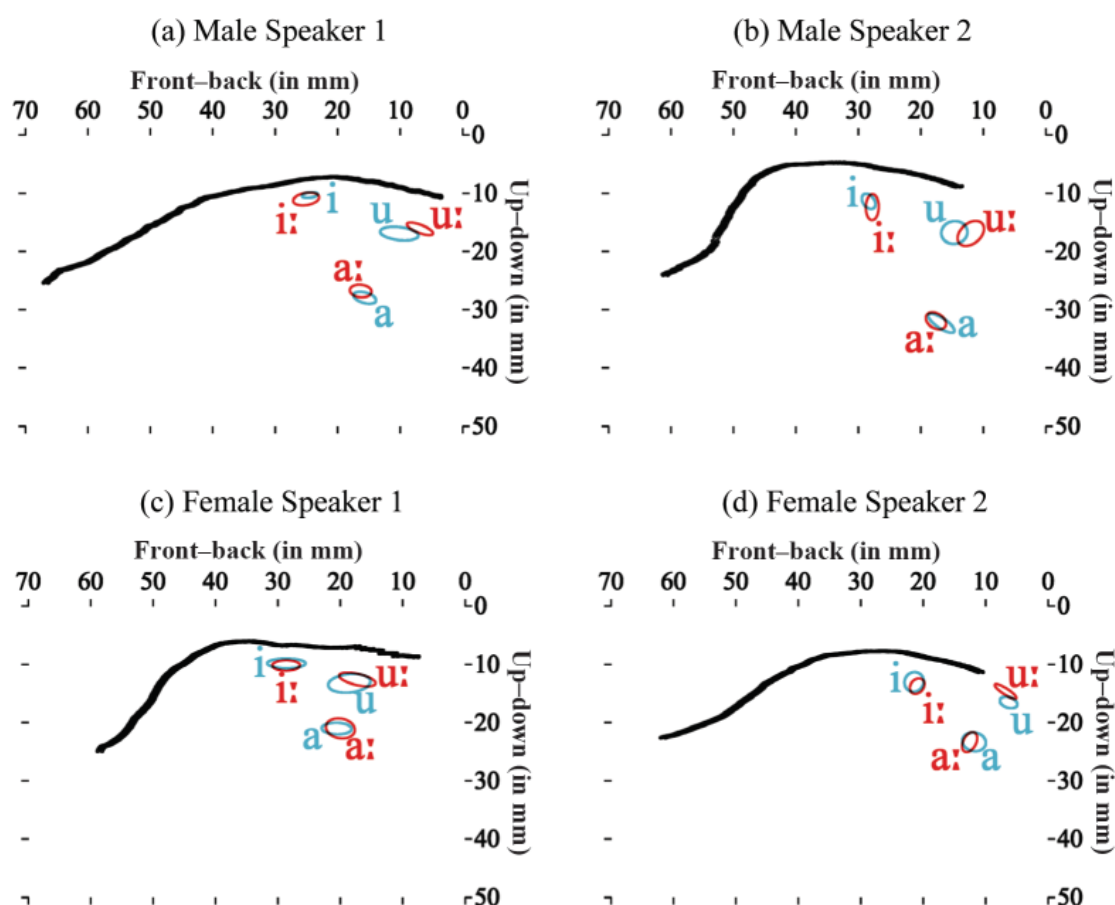


Figures 4a–4d: Tongue contours and tongue positions for the Cantonese shortened vowels [i u a] and the palate contours (thick dark line) for two male and two female speakers

For any vowel, long (Figures 3a–3d) or shortened (Figures 4a–4d), the articulatory data point at the tongue-back position is higher and closer to the opposing palate than the other two points at the tongue-front and tongue-middle positions. This is true for all four speakers. The tongue-back articulatory point is thus taken to be the point of tongue constriction for the vowels.

Figures 5a–5d show the superimposed vowel ellipses for the long vowels [i: u: a:] and shortened vowels [i u a] in the articulatory vowel space for each one of four speakers. Each vowel ellipse was drawn using the function ‘Draw Sigma Ellipses’ provided by the *Praat* program (Boersma & Weenink 2014), based on the 20 tongue constriction points in the *x*- and *y*-dimension for a particular vowel. The two axes of the ellipse are the variations of the ‘*x*, *y*’ positional coordinates centered at the means of the 20 tongue constriction points distributed in the horizontal (*x*) and vertical (*y*) articulatory vowel space. The radii of the ellipse are based on two sigmas or standard deviations representing 86.5% probability of containment for a bivariate normal distribution. Also shown in the figures is the tracing of the mid-sagittal palate contour (thick dark line).

As shown in Figures 5a–5d, the shortened [i] is slightly higher than the long [i:] in the *y*-position. The results of *t*-tests show that the difference in *y*-position between [i:] and [i] is significant for Male Speaker 1 ($p < 0.0001$), Male Speaker 2 ($p < 0.01$), Female Speaker 1 ($p < 0.05$), and Female Speaker 2 ($p < 0.01$). The raise in the *y*-position of the shortened [i] relative to the long



Figures 5a–5d: Superimposed articulatory vowel ellipses for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] and the palate contours (thick dark line) for two male and two female speakers

[i:] appears to constitute a case where the shorter duration of [i] causes the vowel to overshoot articulatorily, rather than undershoot as would be expected to occur. In the articulatory vowel space, [i:] and [i] are similar in the x -position. The difference in the x -position is non-significant for Female Speaker 1 ($p > 0.5$) and Female Speaker 2 ($p > 0.05$), but significant for Male Speaker 1 ($p < 0.05$) and Male Speaker 2 ($p < 0.01$).

Relative to [u:], [u] shifts leftward and downward, which is an indication of the centralization of [u] in the articulatory vowel space. There is a significant difference: (i) in the x -position for Male Speaker 1 ($p < 0.0001$), Male Speaker 2 ($p < 0.0001$), and Female Speaker 1 ($p < 0.05$); and (ii) in the y -position for Male Speaker 1 ($p < 0.0001$), Female Speaker 1 ($p < 0.01$), and Female Speaker 2 ($p < 0.0001$). As [u] in checked CVC syllables is obligatorily followed by [t], the changes in position of the articulatory ellipse for [u] are presumably due to the anticipatory assimilation effect of [t] coupled with the effect of shortened vowel duration, as the effect of [t] alone is expected to cause [u] to undergo fronting, rather than centralization. Furthermore, a comparison of the positions of the tongue-front, or the tongue tip, during [u:] (Figures 3a–3d) and [u] (Figures 4a–4d) shows there is a small forward shift for [u] relative to [u:] for Male Speaker 1

(Figures 3a and 4a), Male Speaker 2 (Figures 3b and 4b), and Female Speaker 1 (Figures 3c and 4c). This corresponds to the fronting of the tongue-back position for [u] for the three speakers. In view of the fact that (i) the lowering of the tongue-back position for [u] is also observed in Male Speaker 1 (Figure 5a) and Female Speaker 1 (Figure 5c), but not Male Speaker 2 (Figure 5b), and (ii) no fronting of the tongue-front or tongue-back position but the lowering of the tongue-back position is observed in Female Speaker 2 (Figure 5d), it may be suggested that the fronting in the centralized [u] is mainly due to the effect of articulatory assimilation of the post-vocalic [t], whereas the lowering is primarily attributed to the effect of shortened vowel duration.

Between [a:] and [a], there is an extensive overlap in the articulatory vowel space for all speakers, which suggests a minimal difference in tongue position. The results of *t*-tests show a significant difference in the *x*-position ($p < 0.01$) or *y*-position ($p < 0.01$) for Male Speaker 1, but only in the *x*-position for Male Speaker 2 ($p < 0.01$) and Female Speaker 2 ($p < 0.01$). As for Female Speaker 1, there is no significant difference in either the *x*-position ($p > 0.1$) or *y*-position ($p > 1.0$).

Repeated measures analysis of variance (ANOVAs) were performed on the three pairs of long and shortened vowels, [i: i], [u: u], and [a: a], for the effects of vowel type and vowel length on the shifts in the *x*- and *y*-positions which are normalized with the *z*-scores. Results show that the vowel type ($p < 0.001$), the vowel length ($p < 0.001$), and the interaction between the vowel type and vowel length ($p < 0.001$) have a significant effect on the shifts in tongue displacement in the *x*- or *y*-dimension. A significant difference in tongue displacement is also observed between any two pairs of long and shortened vowels. This is true for all four speakers, except for the difference between the pairs of [i: i] and [a: a] in the shift in the *x*-position for Female Speaker 1 and the difference between the pairs of [u: u] and [a: a] in the shift in the *y*-position for Male Speaker 1. The different patterns of the shift in the articulatory space for [i u a], relative to [i: u: a:], indicate that the effect of shortened vowel duration is non-uniform across the different vowel types.

In Figures 5a–5d, the orientation of any articulatory vowel ellipse indicates the direction of variances with respect to the tongue constriction location in the *x*-dimension and tongue constriction degree in the *y*-dimension along the vocal tract length during a particular vowel. Tables 2a–2d present the variations in the *x*-position and *y*-position of the tongue constriction during each of the long vowels [i: u: a:] and shortened vowels [i u a] for four speakers. The variations in the *x*-position and *y*-position of the tongue constriction are expressed as the coefficient of variance, that is, the ratio of the standard deviation to the mean of 20 repetitions of each vowel. Also presented in the tables are the means and standard deviations of the *x*-position and *y*-position for the vowels.

As can be seen from the data presented in Tables 2a–2d, in a majority of cases, the variation is smaller in the *y*-position than in the *x*-position. This suggests a smaller degree of articulatory perturbation in tongue constriction degree than in tongue constriction location. The finding is similar to what was reported in Beckman et al. (1995) on the production of the English vowels [i] and [u]. In Cantonese, the exceptions are the long vowel [i:] for all four speakers (Tables 2a–2d) and the shortened [i] for Male Speaker 2 (Table 2b) and Female Speaker 2 (Table 2d). For both vowels, the variation is larger in the *y*-position than in the *x*-position.

Tables 2a–2d: Means ($n = 20$), standard deviations (SD), and ratios of SD to the mean for the x -position and y -position of the tongue constriction for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

(a) Male Speaker 1							
		i:	u:	a:	i	u	a
x -position	Mean	25.05	6.85	16.32	24.44	10.16	15.66
	SD	1.04	1.05	0.86	0.63	1.53	0.89
	SD/Mean	0.041	0.153	0.053	0.026	0.151	0.057
y -position	Mean	10.99	16.13	26.77	10.30	16.94	27.94
	SD	0.56	0.54	0.55	0.24	0.58	0.51
	SD/Mean	0.051	0.033	0.021	0.023	0.034	0.018

(b) Male Speaker 2							
		i:	u:	a:	i	u	a
x -position	Mean	27.77	12.02	17.60	28.34	14.65	16.74
	SD	0.52	1.06	0.80	0.55	1.05	1.04
	SD/Mean	0.019	0.088	0.046	0.019	0.072	0.062
y -position	Mean	12.40	16.92	31.89	14.41	16.76	32.29
	SD	1.12	1.07	0.73	0.66	0.97	0.81
	SD/Mean	0.090	0.063	0.023	0.058	0.058	0.025

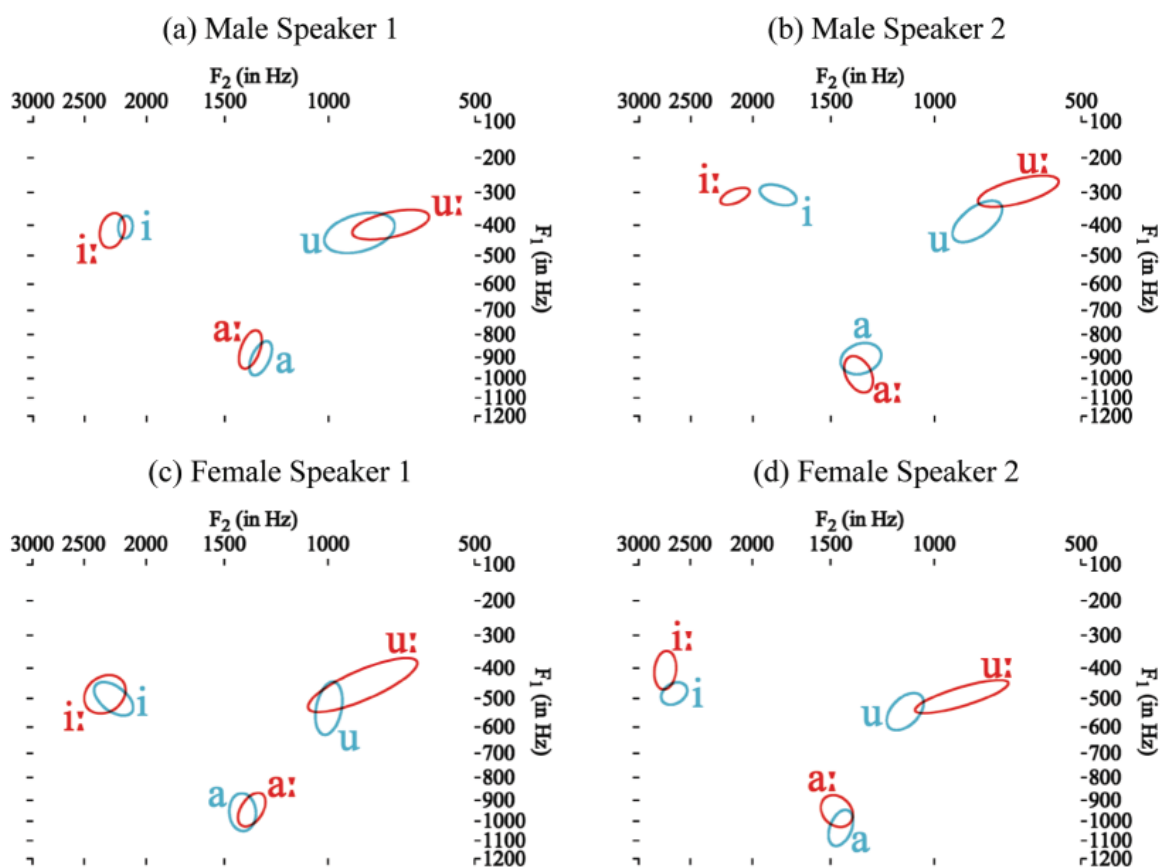
(c) Female Speaker 1							
		i:	u:	a:	i	u	a
x -position	Mean	28.68	17.26	19.98	28.64	18.55	20.57
	SD	1.10	1.46	1.15	1.53	1.69	1.19
	SD/Mean	0.038	0.084	0.058	0.053	0.091	0.058
y -position	Mean	10.16	12.58	21.00	9.82	13.25	21.02
	SD	0.44	0.60	0.86	0.42	0.82	0.050
	SD/Mean	0.043	0.048	0.041	0.043	0.062	0.024

(d) Female Speaker 2							
		i:	u:	a:	i	u	a
x -position	Mean	20.98	6.82	12.57	21.47	6.33	11.86
	SD	0.59	0.87	0.60	0.77	0.70	0.94
	SD/Mean	0.028	0.127	0.048	0.036	0.110	0.079
y -position	Mean	13.74	14.67	23.35	13.01	16.56	23.31
	SD	0.67	0.64	0.86	0.86	0.50	0.81
	SD/Mean	0.049	0.043	0.037	0.066	0.030	0.035

3.3 Spectral characteristics

Figures 6a–6d show the superimposed acoustical vowel ellipses for the Cantonese long vowels [i:, u:, a:] and shortened vowels [i u a] in the F_1 – F_2 plane for four speakers. The figures are on the Bark scale, with the origin at the top right corner. Each of the acoustical vowel ellipses was drawn using the *Praat* program (Boersma & Weenink 2014) and based on the radii of two standard deviations along 20 F_1 F_2 vowel formant data points for a particular vowel.

A comparison of Figures 6a–6d and Figures 5a–5d for any one of the speakers shows that the degree of overlap between the long vowels and the shortened equivalents is smaller in the acoustical vowel space than in the articulatory vowel space. The degree of overlap of the acoustical vowel ellipses is particularly small: (i) between [a:] and [a] for Male Speaker 1 (Figure 6a); and (ii) between [i:] and [i] and between [u:] and [u] for Male Speaker 2 (Figure 6b) and Female Speaker 2 (Figure 6d).



Figures 6a–6d: Superimposed acoustical vowel ellipses in the F_1 – F_2 plane for the Cantonese long vowels [i:, u:, a:] and shortened vowels [i u a] for two male and two female speakers

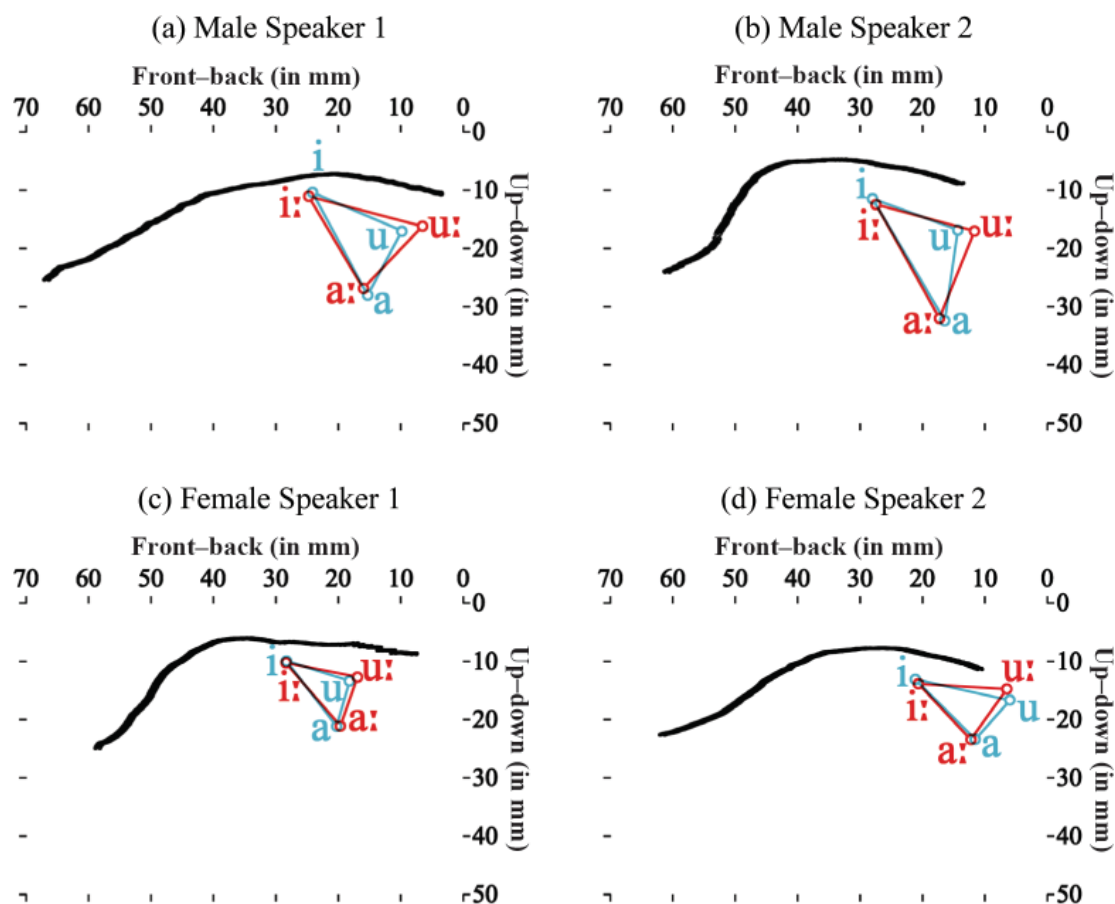
For all speakers, the overlap between the vowel ellipses in the acoustical vowel space for [u:] and [u] is only partial (Figures 6a–6d), with [u] being positioned to the lower left of [u:] due to an increase in F_1 and F_2 . The results of *t*-tests show that the differences in F_1 and F_2 between [u:] and [u] are significant for Male Speaker 1 ($p < 0.01$; $p < 0.0001$), Male Speaker 2 ($p < 0.0001$; $p < 0.0001$), Female Speaker 1 ($p < 0.0001$; $p < 0.0001$), and Female Speaker 2 ($p < 0.0001$). The data indicate that the effect of the shortened duration coupled with the anticipatory assimilation effect of the post-vocalic [t] cause [u] to undergo centralization. As shown in Figures 5a–5d, the centralization of [u] relative to [u:] is also observed in the articulatory vowel space for the speakers.

For all four speakers, [i] relative to [i:] undergoes formant undershoot as indicated by a decrease in F_2 (Figures 6a–6d). The results of *t*-tests show that the difference in F_2 between [i:] and [i] is significant for Male Speaker 1 ($p < 0.0001$), Male Speaker 2 ($p < 0.0001$), Female Speaker 1 ($p < 0.01$), and Female Speaker 2 ($p < 0.001$). The difference in F_1 between [i:] and [i] is significant only for Female Speaker 2 ($p < 0.0001$). As shown in Figures 5a–5d, for all four speakers, [i] relative to [i:] does not undergo target reduction; instead it shifts slightly upward in the articulatory vowel space. It is assumed that modification of the vocal tract shape, aside from the tongue displacement, may have been involved during the vowel [i], causing formant undershoot or vowel centralization.

The shortened [a], relative to the long [a:], in general does not undergo centralization in the acoustical vowel space (Figures 6a, 6c, and 6d), except for Male Speaker 2 where the vowel is slightly higher than the long equivalent [a:] (Figure 6b). As shown in Figures 5a–5d, for all four speakers, the difference in position between [a] and [a:] in the articulatory vowel space is small. Unlike [i:] and [i] and also [u:] and [u], there is a lack of pattern for the difference in formants or tongue position between [a:] and [a]. Thus, the effect of the shortened vowel duration on the formants or tongue position varies across the vowel types.

3.4 Articulatory and acoustical vowel space

Figures 7a–7d show the superimposed vowel loops for the two sets of Cantonese vowels, the long [i: u: a:] and shortened [i u a], in the articulatory vowel space for four speakers. The thick dark line shows the mid-sagittal palate contour. The articulatory vowel loops were drawn by connecting the tongue constriction points for the three vowels of each set, with each point representing the mean of 20 repetitions of the position in the *x*- and *y*-dimensions for one of the three vowels of the set.

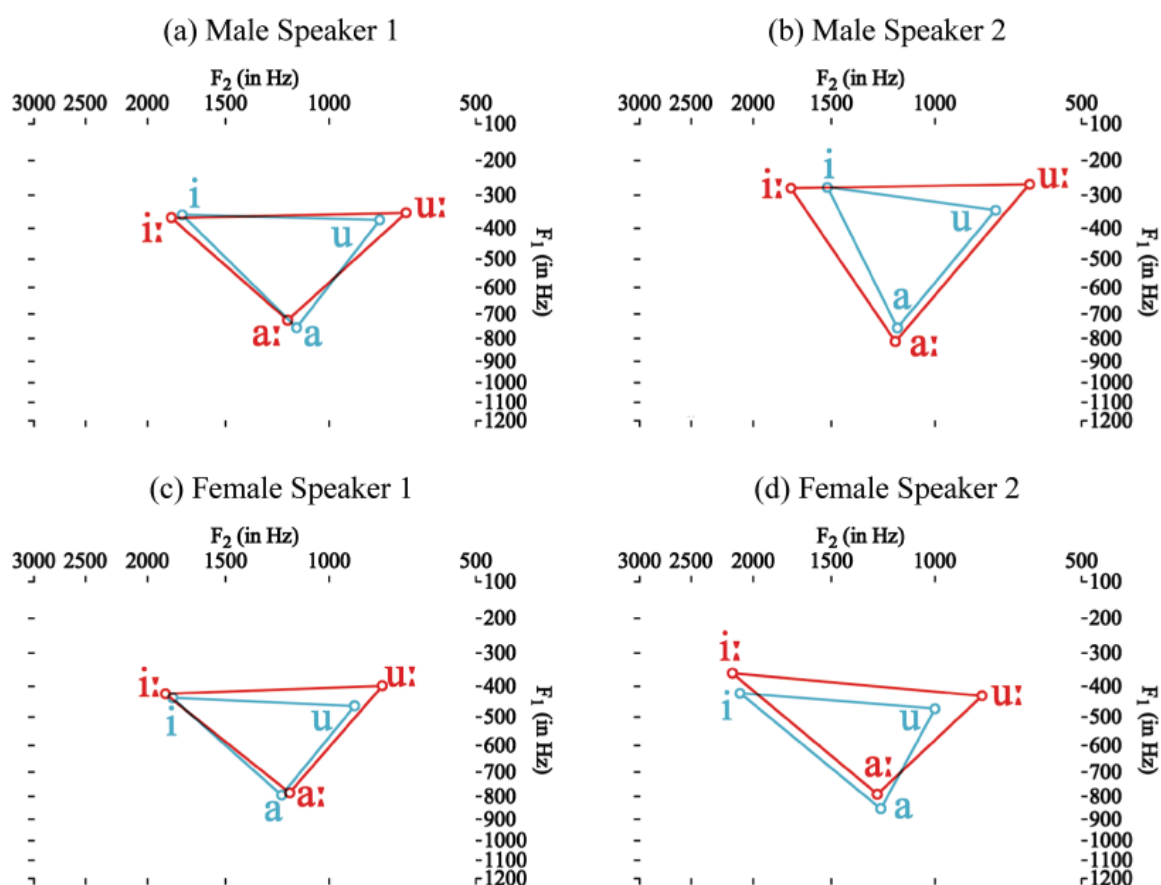


Figures 7a–7d: Superimposed articulatory vowel loops for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] and the palate contours (thick dark line) for two male and two female speakers

Figures 8a–8d show the superimposed vowel loops for the two sets of vowels in the acoustical vowel space. The acoustical vowel loops were drawn by connecting the data points of the three vowels of each set, with each point representing the mean F_1 and F_2 for 20 repetitions of one of the three vowels of the set.

A comparison of the two sets of the articulatory vowel loops in Figures 7a–7d shows a reduction in the vowel loop area for the shortened vowels [i u a] relative to the long vowels [i: u: a:]. This is true for four speakers, with a smaller reduction for Female Speakers 1 and 2 (Figures 7c and 7d) than Male Speakers 1 and 2 (Figures 7a and 7b). Similarly, there is also a reduction in the acoustical vowel loop area for [i u a] relative to [i: u: a:] for the four speakers (Figures 8a–8d).

By applying Heron's formula (Weisstein 2014) for calculating the area of an irregular polygonal shape (Jacewicz et al. 2007; Neel 2008), both the articulatory and acoustic vowel loop areas for the two sets of vowels were obtained. The area data presented in Table 3 for the two sets of vowels show that for all four speakers there is a reduction in the areas of both the articulatory and acoustical vowel loops for [i u a] relative to [i: u: a:]. The reduction in the acoustical vowel loop area is



Figures 8a–8d: Superimposed acoustical vowel loops in the F_1 – F_2 plane for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

due to the undershoot of F_1 and/or F_2 for both [i] and [u] in the acoustical vowel space, whereas the reduction in the articulatory vowel loop is primarily due to the vowel target reduction of [u] in the articulatory vowel space.

Table 3: Areas of the articulatory vowel loops (in mm^2) and acoustical vowel loops (on Bark^2) for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

	Articulatory vowel loop		Acoustical vowel loop	
	[i: u: a:]	[i u a]	[i: u: a:]	[i u a]
Male Speaker 1	121.15	96.78	11.88	10.46
Male Speaker 2	130.49	111.94	17.87	10.66
Female Speaker 1	51.36	42.68	10.87	8.12
Female Speaker 2	64.19	60.91	13.01	9.83

3.5 Articulatory–acoustical relationship

Beckman et al. (1995) reported that in the English vowels [i] and [u], formant frequencies are less sensitive to variation in tongue constriction location in the x -dimension than variation in tongue constriction degree in the y -dimension. To test whether this is also true of the Cantonese vowels, the R^2 values were obtained to establish the correlations between F_1 and variation in tongue height, that is, the tongue constriction degree in the y -dimension, and between F_2 and variation in backness of the tongue, that is the tongue constriction location in the x -dimension. Table 4 presents two types of R^2 values, ranging from zero to one, for each of the Cantonese vowels [i: u: a:] and [i u a], with R^2_{F1-y} representing the square of the correlation coefficient of F_1 and variation in the y -position of tongue constriction, and R^2_{F2-x} representing the square of the correlation coefficient of F_2 and variation in the x -position of tongue constriction. A larger R^2_{F1-y} or R^2_{F2-x} is taken to indicate a higher level of sensitivity of formant frequencies to variation in vertical or horizontal tongue displacement.

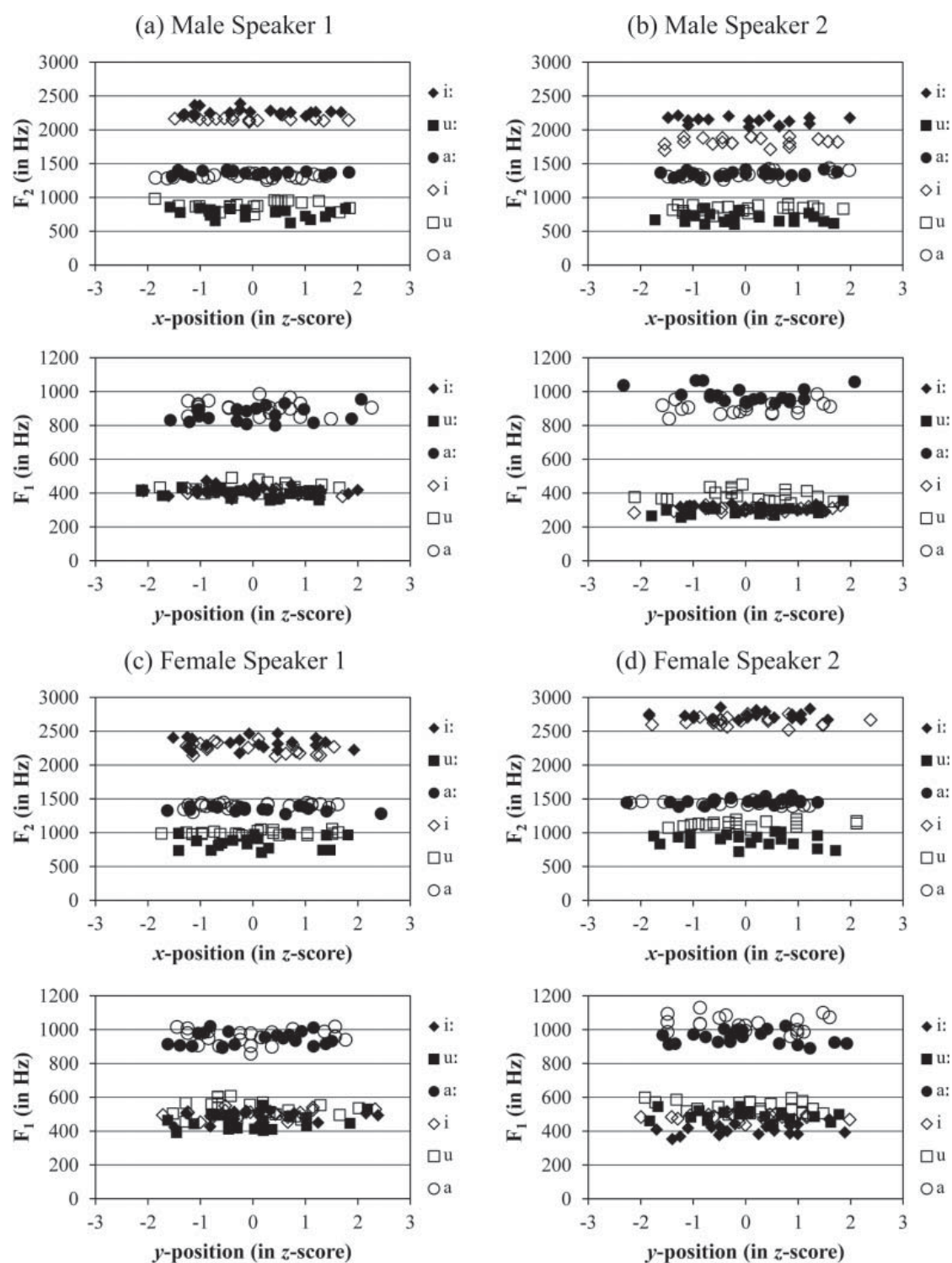
Table 4: R^2_{F1-y} (the square of the correlation coefficient of F_1 and variation in the y -position of tongue constriction), and R^2_{F2-x} (the square of the correlation coefficient of F_2 and variation in the x -position of tongue constriction) for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

	Male Speaker 1		Male Speaker 2		Female Speaker 1		Female Speaker 2	
	R^2_{F1-y}	R^2_{F2-x}	R^2_{F1-y}	R^2_{F2-x}	R^2_{F1-y}	R^2_{F2-x}	R^2_{F1-y}	R^2_{F2-x}
i:	0.0407	0.0187	0.0800	0.0614	0.0491	0.0073	0.1060	0.0005
u:	0.1017	0.0868	0.2130	0.0217	0.0725	0.0005	0.0574	0.0560
a:	0.0884	0.0013	0.0558	0.0751	0.0184	0.1574	0.0309	0.1530
i	0.1155	0.1215	0.0783	0.0147	0.0560	0.1078	0.0050	0.0012
u	0.0185	0.0002	0.0145	0.0604	0.0572	0.0503	0.1020	0.1404
a	0.0375	0.0090	0.0296	0.3510	0.0015	0.1186	0.0348	0.0451

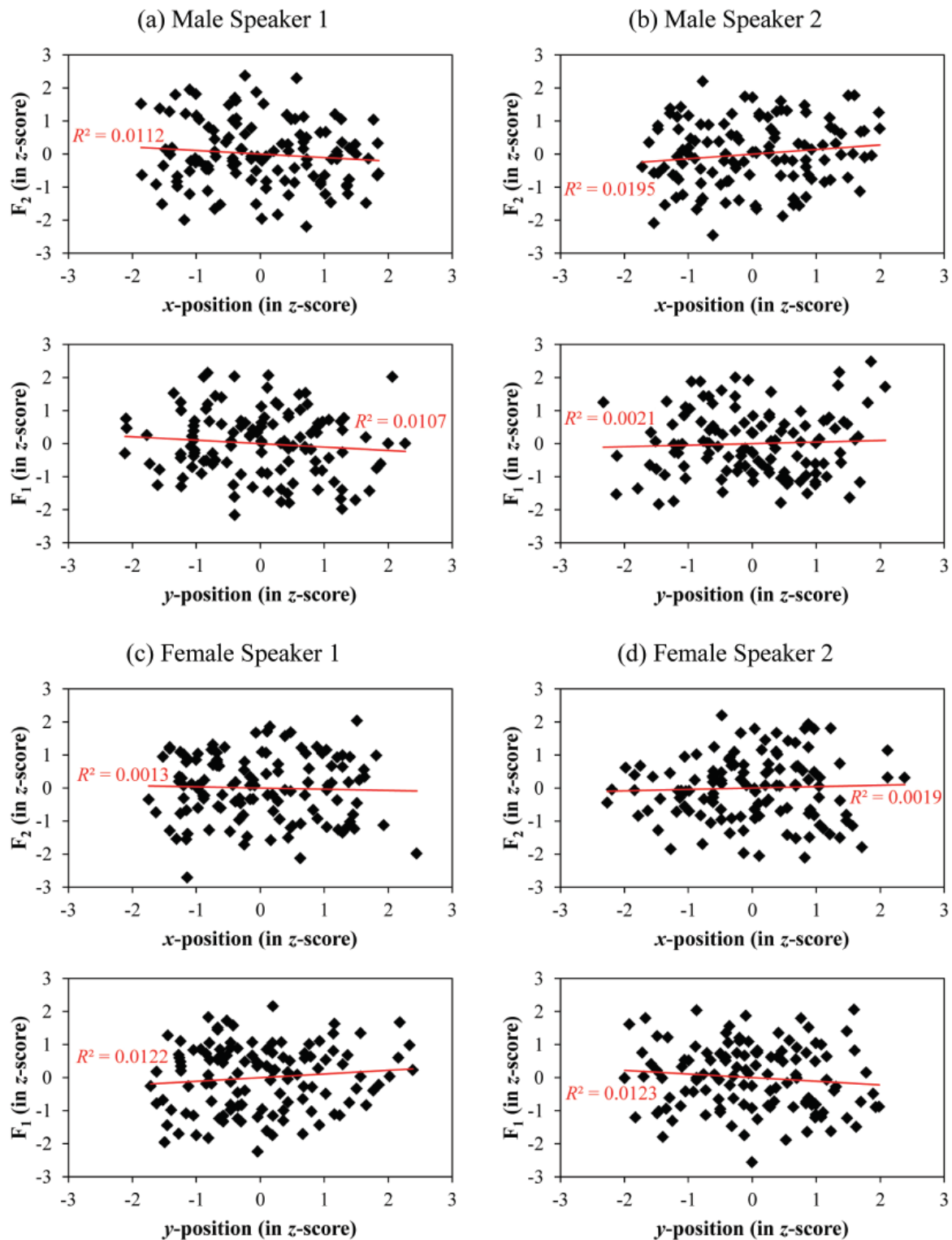
As can be seen in the data presented in Table 4, R^2_{F1-y} and R^2_{F2-x} for any vowel are small. For Male Speaker 1, they are close to 0.1 or near zero in all cases. This is also true for Female Speakers 1 and 2. As for Male Speaker 2, R^2_{F1-y} and R^2_{F2-x} are smaller than 0.1 in nearly all cases, except the R^2_{F1-y} for [u:] (0.2130) and R^2_{F2-x} for [a] (0.3510). The near zero R^2_{F1-y} and R^2_{F2-x} values indicate that the levels of sensitivity of F_1 and F_2 to the variations in tongue displacement in the x - and y -dimension are low in each of the vowels [i: u: a:] and [i u a].

A comparison of R^2_{F1-y} and R^2_{F2-x} for each of the Cantonese vowels [i: u: a:] and [i u a] presented in Table 4 shows that for all four speakers R^2_{F1-y} is larger than R^2_{F2-x} for the long vowels [i:] and [u:]. For the other vowels, this is true only for some speakers—for instance: (i) the long [a:] for Male Speaker 1; (ii) the shortened [i] for Male Speaker 2 and Female Speaker 2; (iii) the shortened [u] for Male Speaker 1 and Female Speaker 1; and (iv) the shortened [a] for Male Speaker 1. That the formant frequencies are less sensitive to variation in tongue constriction location in the x -dimension than variation in tongue constriction degree in the y -dimension is true of the long vowels [i:] and [u:], but not for the long [a:] and the shortened [i u a].

Figures 9a–9d show the scatter plots of F_1 against the variation in the y -position of tongue constriction (lower panels) and F_2 against the variation in the x -position (upper panels) for each



Figures 9a–9d: Scatter plots of F_1 against the variation in the y-position of the tongue constriction (in z-score) in the lower panels and F_2 against the variation in the x-position of the tongue constriction (in z-score) in the upper panels for each of the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers



Figures 10a–10d: Scatter plots of F_1 (in z-score) against the y -position of the tongue constriction (in z-score) in the lower panels and F_2 against the x -position of the tongue constriction (in z-score) in the upper panels for the Cantonese long vowels [i: u: a:] and shortened vowels [i u a] for two male and two female speakers

of the six vowels for four speakers, displaying the linearity between the formant values and the displacement of tongue constriction. In the figures, the variations in the x - and y -positions of tongue constriction for the vowels are normalized with the z -scores, showing the variation in tongue displacement from the mean position at zero. As shown in the figures, the scattered data points, representing F_1 against the variation in the y -position of tongue constriction and F_2 against the variation in the x -position of tongue constriction, for each vowel are flatly distributed along the horizontal axis. This not only indicates a low sensitivity of the formant values to the variations in tongue displacement, but also denotes a lack of linear relationship between the two variables in each vowel.

Figures 10a–10d show the scatter plots for the z -scores of F_1 against the variation in the y -position of tongue constriction (lower panels) and F_2 against the variation in the x -position (upper panels) for all the six vowels [i: u: a:] and [i u a] for four speakers, exhibiting the variations from the mean value or from the mean position at zero. Regression lines and R^2 values (the square of the correlation coefficient of the two variables) are also given in the figures.

As can be seen in Figures 10a–10d, for all four speakers the regression line is flat along the horizontal axis or only slightly sloped, and the R^2 value is near zero. It follows that the relationships between the formant values and variations in tongue displacement, that is, between F_1 and variation in the y -position of tongue constriction and between F_2 and variation in the x -position of tongue constriction, are not linear for the six Cantonese point vowels.

4. Discussion

4.1 Vowel target undershoot

Lindblom (1963), in an investigation of the effect of duration on the formant frequencies of the Swedish vowels in nonsense CVC syllables, held that short duration is the cause of vowel reduction. As the vowel shortens, there is insufficient time for articulators to reach the vowel target, resulting in undershoot of formant frequencies in the acoustical vowel space, relative to the target F -pattern associated with the long vowel. As presented in the foregoing sections, in Cantonese the F -patterns of the shortened vowels [i u a] differ from those of the long equivalents [i: u: a:]. There is a marked increase in F_1 and F_2 for [u] relative to [u:], indicating vowel formant centralization in [u]. As pointed out earlier, since [u] is obligatorily followed by a front consonant [t] within the test syllable, the centralization of the shortened [u] in the acoustical vowel space is presumably due to the effect of anticipatory assimilation of the post-vocalic consonant, as well as the effect of shortened vowel duration, as the effect of [t] alone is expected to cause [u] to undergo fronting, rather than centralization. For [i], formant undershoot, relative to [i:], takes place in F_2 , but not F_1 . As for [a], no substantial change in F_1 and F_2 relative to [a:] is observed. Based on formant data on the Cantonese shortened vowels [i u a], it may be concluded that the effect of duration on the formant frequencies of the Cantonese point vowels is not wholly predicted by Lindblom's (1963) theory of vowel target undershoot.

For the Cantonese shortened vowel [u], formant undershoot in the acoustical vowel space is paralleled by vowel centralization in the articulatory vowel space. This is not the case for the

Cantonese shortened vowel [i], as the formant undershoot in the acoustical vowel space is not accompanied by vowel target reduction in the articulatory vowel space. That formant undershoot is not always accompanied by vowel target reduction was also reported in Mooshammer & Geng's (2008) articulatory–acoustical investigation of German vowels. In German, the back vowels undergo target undershoot in both the acoustical and articulatory vowel space, whereas for the front vowels, formant undershoot in the acoustical vowel space is not paralleled by vowel centralization in the articulatory vowel space. In theory, a change in articulatory position results in a change in the F-pattern, and the reverse is expected to be true. That there is no change in the articulatory vowel space paralleling a change in the acoustical vowel space for the front vowels in Cantonese and German may be because there are other articulatory modifications in the mouth cavity during the vowels other than a change in tongue position.

In Cantonese, unlike the shortened high vowels [i] and [u], the extent of formant undershoot or articulatory vowel reduction is much smaller for the shortened low vowel [a]. Contrary to the Cantonese vowel [a], the Creek short vowel [a] undergoes extensive formant undershoot (Johnson & Martin 2001). Similarly, in German, the unstressed low vowels with a reduced duration undergo substantial vowel target reduction and formant undershoot (Mooshammer & Geng 2008). The cross-language difference may be due to: (i) a difference in vowel duration, with the Cantonese shortened vowel [a] (170 ms) being longer than the short vowel [a] in Creek (124 ms) and the unstressed low vowels in German (30–50 ms); and (ii) the co-articulation effect of the neighboring segments, as the Creek vowel [a] occurs in test polysyllabic words (Johnson & Martin 2001) and the German low vowels in test disyllabic words (Mooshammer & Geng 2008). In the case of the Cantonese shortened vowel [a], it occurs in a test monosyllabic word.

For both Cantonese and German vowels, the shortening of vowel duration results in a reduction in the acoustical and articulatory vowel space area. In both cases, the reduction in the acoustical vowel space results from the formant undershoot of the front and back vowels, whereas the reduction in the articulatory vowel space is attributed primarily to the centralization of the back vowels. It may be concluded that: (i) the effect of vowel shortening on tongue displacement is larger in the back vowels than the front ones, although in the case of the Cantonese shortened [u] further investigation is needed, as the vowel is presumed to be concurrently affected by the post-vocalic consonant [t] and the shortened vowel duration, and (ii) the articulatory–acoustical relationships in vowel production are non-uniform across the vowel types.

4.2 Variability in tongue displacement and articulatory–acoustical relationship

Beckman et al. (1995), in an investigation of articulatory and acoustic variability in the English vowels, including the point vowels [i], [u], and [ɑ], in CVC syllables, reported that their findings provide support for quantal theory (Stevens 1972, 1989). As stated in Beckman et al. (1995):

Quantal Theory is that speech sounds are preferred cross-linguistically if their articulatory targets take advantage of regions and dimensions where acoustic parameters are relatively insensitive to articulatory variation. Cross-linguistically, it is the case that among vowels, [i], [ɑ], and [u] are the most preferred. (p. 471)

The point vowels [i u ɑ] with a narrow constriction along the vocal tract are said to be quantal due to their formant frequencies being relatively insensitive to articulatory variation. The study showed that in English: (i) the primary variability in tongue displacement for the vowels [i] and [u] is in the tongue constriction location in the *x*-dimension, rather than the tongue constriction degree in the *y*-dimension; (ii) the formant frequencies for [i] and [u] are more stable and less sensitive to the variation in tongue constriction location in the *x*-dimension than the variation in tongue constriction degree in the *y*-dimension; and (iii) the articulatory–acoustical relationship in [i] and [u] is considered to be quantal. In Cantonese: (i) smaller variability in the tongue constriction degree in the *y*-dimension than the tongue constriction location in the *x*-dimension is observed for the high back vowels [u: u], but not the high front [i: i]; and (ii) for the long vowels [i:] and [u:], but not the shortened [i] and [u], the formant frequencies are less sensitive to variation in tongue constriction location in the *x*-dimension than variation in tongue constriction degree in the *y*-dimension. The data suggest that Cantonese high back point vowels are less quantal relative to the high front equivalents, as are the Cantonese shortened point vowels relative to the long equivalents. Beckman et al. (1995) also reported that there is a quantal relationship between articulation and formants in the English high vowels [i] and [u], but not the low vowel [ɑ]. In Cantonese, the quantal relationship is found in the long vowels [i:] and [u:], but not the long vowel [a:] and the shortened vowels [i u ɑ]. Thus, there are similarities and differences in articulatory–acoustical relationships for the point vowels between the two languages.

5. Concluding remarks

The paper has presented: (i) the effect of duration on the articulation and formant frequencies of two sets of Cantonese point vowels, long and shortened; and (ii) the articulatory–acoustical relationships in the production of these vowels. As the effect of shortened vowel duration and the articulatory–acoustical relationships in vowel production vary in different languages, data should be collected from more languages in order to reach a cross language generalization where duration affects articulation and formant frequencies of vowels. And, as the Cantonese shortened vowel [u] is concurrently affected by the post-vocalic consonant [t] and the shortened vowel duration, further study is required in order to investigate the interplay between the two factors and to determine the contribution of each factor to vowel reduction in Cantonese.

References

- Beckman, Mary E., Tzyy-Ping Jung, Sook-hyang Lee, Kenneth de Jong, Ashok K. Krishnamurthy, Stanley C. Ahalt, K. Bretonnel Cohen, & Michael J. Collins. 1995. Variability in the production of quantal vowels revisited. *Journal of the Acoustical Society of America* 97.1:471–490.
- Behne, Dawn, Bente Moxness, & Anne Nyland. 1996. Acoustic-phonetic evidence of vowel quantity and quality in Norwegian. *Speech, Music and Hearing, Quarterly Progress and Status Report (TMH-QPSR)* 37.2:13–16.
- Boersma, Paul, & David Weenink. 2014. Praat: doing phonetics by computer. Accessed online 4 July 2014 at <http://www.fon.hum.uva.nl/praat/>.

- Gracco, Vincent L. 1995. Electromagnetic articulography: a brief overview. *Proceedings of the XIIIth International Congress of Phonetic Sciences (ICPhS 95)*, ed. by Kjell Elenius & Peter Branderud, 58–61. Stockholm: Congress organisers at KTH and Stockholm University.
- Hoole, Philip. 1996. Issues in the acquisition, processing, reduction and parameterization of articulo-graphic data. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM)* 34:158–173.
- Hoole, Philip. 1999a. On the lingual organization of the German vowel system. *Journal of the Acoustical Society of America* 106.2:1020–1032.
- Hoole, Philip. 1999b. Articulatory–acoustic relations in German vowels. *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS 99)*, ed. by John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey, 2153–2156. San Francisco: Linguistics Department, University of California.
- Jacewicz, Ewa, Robert A. Fox, & Joseph Salmons. 2007. Vowel space areas across dialects and gender. *Proceedings of the XVIth International Congress of Phonetic Sciences (ICPhS 2007)*, ed. by Jürgen Trouvain & William J. Barry, 1465–1468. Saarbrücken: Saarland University.
- Johnson, Keith, & Jack Martin. 2001. Acoustic vowel reduction in Creek: effects of distinctive length and position in the word. *Phonetica* 58:81–102.
- Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35.11:1773–1781.
- Mooshammer, Christine, & Christian Geng. 2008. Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association* 38.2:117–136.
- Neel, Amy T. 2008. Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research* 51.3:574–585.
- Perkell, Joseph S., & Marc H. Cohen. 1989. An indirect test of the quantal nature of speech in the production of the vowels /i/, /a/ and /u/. *Journal of Phonetics* 17:123–133.
- Perkell, Joseph S., Marc H. Cohen, Mario A. Svirsky, Melanie L. Matthies, Iñaki Garabieta, & Michel T.T. Jackson. 1992. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America* 92.6:3078–3096.
- Perkell, Joseph S., & Winston L. Nelson. 1985. Variability in production of the vowels /i/ and /a/. *Journal of the Acoustical Society of America* 77.5:1889–1895.
- Stevens, Kenneth N. 1972. The quantal nature of speech: evidence from articulatory–acoustic data. *Human Communication: A Unified View*, ed. by Edward E. David, Jr. & Peter B. Denes, 51–66. New York: McGraw-Hill.
- Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17:3–45.
- Tabain, Marija. 2003a. Effects of prosodic boundary on /aC/ sequences: acoustic results. *Journal of the Acoustical Society of America* 113.1:516–531.
- Tabain, Marija. 2003b. Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of the Acoustical Society of America* 113.5:2834–2849.
- Tabain, Marija, & Pascal Perrier. 2005. Articulation and acoustics of /i/ in pre-boundary position in French. *Journal of Phonetics* 33:77–100.
- Tabain, Marija, & Pascal Perrier. 2007. An articulatory and acoustic study of /u/ in pre-boundary position in French: the interaction of compensatory articulation, neutralization avoidance and featural enhancement. *Journal of Phonetics* 35:135–161.

- Weisstein, Eric W. 2014. Heron's formula. From MathWorld—A Wolfram Web Resource. Accessed online 4 July 2014 at <http://mathworld.wolfram.com/HéronsFormula.html>.
- Zee, Eric. 1999. Chinese (Hong Kong Cantonese). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, ed. by the International Phonetic Association, 58–60. Cambridge: Cambridge University Press.

[Received 12 June 2014; revised 31 December 2014; accepted 13 April 2015]

Department of Linguistics and Translation
City University of Hong Kong
83 Tat Chee Avenue
Kowloon Tong, Hong Kong
w.s.lee@cityu.edu.hk

廣州話元音發音與聲學之關係

李蕙心

香港城市大學

本文分析廣州話兩組邊緣元音 [i: u: a:] 和 [i u a] 的發音與聲學之關係。長元音 [i: u: a:] 在 CV 開音節中出現；[i u a] 在 CVC 閉音節中時長縮短。根據兩男和兩女發音人的舌位及共振峰資料顯示，(i) 相對於長高元音 [i: u:]，縮短了的高元音 [i u] 的共振峰在聲學元音區域中不達目標，同時 [u] 在發音元音區域中也不達目標；不過，(ii) 這種時長對高元音 [i u] 的影響，在低元音 [a] 中沒有產生。有關元音的發音與聲學關係方面，(i) [i: u: a:] 和 [i u a] 的 F_1 及 F_2 共振峰，對於舌頭與上顎間收窄點的 x (前後) 位置及 y (上下) 位置移位的敏感度屬低，這表示廣州話的邊緣元音具有 quantal 的性質；及 (ii) 邊緣元音 [i: u:]，而非 [a:] 及 [i u a]，其 F_1 及 F_2 共振峰對於舌頭與上顎間收窄點的 x 位置移位的敏感度，比對 y 位置移位的敏感度為低，這表示 (a) 低邊緣元音的 quantal 性質的程度比高邊緣元音的為低，及 (b) 縮短了的邊緣元音的 quantal 性質的程度比長邊緣元音的為低。本文的分析結果跟相關的語音學理論作討論，如元音目標減縮理論及 quantal 理論，也與英語中邊緣元音的發音與聲學關係作比較。

關鍵詞：廣州話元音，發音與聲學關係，元音目標減縮理論，quantal 理論