An Aerodynamic Analysis of Intonation in Hong Kong Cantonese

Yuk-Man Cheung

Phonetics Lab, Department of Chinese, Translation and Linguistics,
City University of Hong Kong
y.m.cheung@student.cityu.edu.hk

Abstract

This paper presents an aerodynamic analysis of intonation on reading speech in Hong Kong Cantonese. In this study, the variation in oral airflow rate, fundamental frequency (F_0), as well as intensity level of the acoustic signal is studied. The goals of this study are to evaluate the correlations among fundamental frequency of vibration of the vocal folds, rate of oral airflow and intensity level, and to search for (if any) the domain of the 'basic breath unit' within paragraph. The data were collected and analyzed on three subjects who are native speakers of Cantonese. The results suggest first that in spontaneous speech, the variation in oral airflow rate may be used to facilitate some changes in F_0 contour, whereas the correlation between the variation in intensity level and F_0 changes is not significant in this study, and second the 'basic breath unit' at paragraph level does not necessarily match the domain of syntactic units, mainly clauses or sentences, and unfilled pause. The frequency of the occurrences of 'basic breath unit' within paragraph varies across speakers. Such interspeaker differences may due to the differences in speaking style and basic physical need among speakers.

1. Introduction

The physiological correlates of vocal cord vibration and intonation pattern have been extensively studied [1, 2, 8, 9, 10, 12, 13]. On one hand, Ohala [10] and Collier [4] showed that the laryngeal muscles, especially the cricothyroid muscle, participate actively in controlling major F_0 changes. On the other hand, subglottal pressure has been shown to be the dominant factor in controlling the fall in F_0 contour at the end of sentences and the rise in pitch on emphasized words in the investigation of Lieberman [9]. The latter point of Lieberman was supported by Ladefoged [6, 7]. In his early aerodynamic studies of laryngeal activity, the momentary increases in subglottal pressure were shown to be noticeable on emphasized words and an increase in airflow (or subglottal pressure) led to an increase in pitch.

Many experimental studies have investigated the acoustic correlates of tone and intonation in Chinese, Mandarin [3, 11, 14] and Cantonese [15, 16]. However, little work has been done on the aerodynamics of tone language. In Cantonese, there are six lexical tones, [˥, ˧, ˨, ˩, ˥˧, ˧˨] or 55, 33, 22, 21, 25, 23 respectively. It is worthwhile to know more about the aerodynamic aspect of intonation in a tone language, which has a tone system as complicated as Cantonese and involves tone and intonation interaction.

This study examines (1) the correlations among variations in oral airflow rate, variations in intensity level, and F_0 changes at paragraph level in Hong Kong Cantonese, and (2) the domain of 'basic breath unit' in continuous speech according to the data of oral airflow rate. The results of this study are useful information for synthesizing speech unit at paragraph level that involves aerodynamic specification in a tone language.

2. Methods

2.1. Materials and Subjects

The production material in this study was the spoken text of the fable 'The North Wind and the Sun' in vernacular Hong Kong Cantonese [5]. The spoken text was produced by three female native speakers in spontaneous speech. All of the subjects were native speakers of Hong Kong Cantonese who were in twenties. They were asked to utter the whole text at a normal rate of speech.

2.2. Measurements

The recordings were carried out in a sound-proof booth (IAC). The Scicon PCquirer multi-channel data acquisition system, fitted with an oral mask and a microphone, was used to record the data. The speakers were wearing the mask at the time they uttered the text. The speech materials were analyzed, using the Scicon PCquirer 6.0 software. The sampling rate for capturing the audio signals was at a default 11,000 Hz. The DC channel for the rate of oral airflow was at a default 1,375 Hz. The audio signals together with the data of oral airflow rate were simultaneously transferred to the interface system through the audio and airflow transducers and subsequently filtered, digitized and then stored into the computer. The F_0 contours and intensity curves of the audio signals were obtained, using the pitch and intensity synchronous method provided by the software. The values of F_0 contour and intensity for every 10ms were measured and logged. The logged data were used to compute the correlation coefficient between F_0 and intensity. The display of the waveform, rate of oral airflow (in ml/sec), intensity (in dB) and F_0 contour (in Hz) for a monosyllable is shown in Fig. 1.

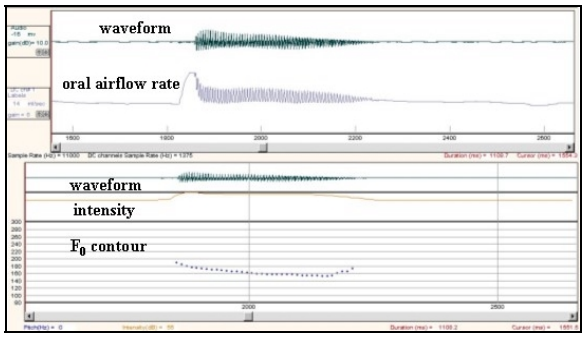

Figure 1: *The display of waveform, rate of oral airflow, intensity and F_0 contour for a monosyllable.*

3. Results

3.1. Oral airflow rate, F_0 and intensity

With a view to discussing the correlations among the oral airflow rate, intensity and F_0 contour at paragraph level, we could get some ideas from such correlations at syllable level. Fig. 2 displays the oral airflow rate, intensity and F_0 contour for the six lexical tones in Cantonese associated with a meaningful monosyllable [si] produced by a female native speaker. It can be seen that without the influence of tone and intonation interaction, the correlation between the pattern of oral airflow rate variation and F_0 contour in monosyllable is quite well. That is, the rate of oral airflow rises for the rising F_0 contour, as in [si25] 史 'history' and [si23] 市 'city' ; falls for the falling contour, as in [si21] 時 'time' and be stable for the level F_0 contour, as in [si55] 思 'to ponder' , [si33] 試 'to try' and [si22] 事 'matter' . It should be noted that although the pattern of oral airflow rate and F_0 changes is similar, the slope and extent of variation in F_0 contour is independent of that in oral airflow rate. The data of the three level tones in Cantonese shows that the difference in the value of F_0 contour among the three tones is not controlled by the variation in oral airflow rate. According to the data in Fig. 2, there seems to be no correlation between intensity and F_0, as well as intensity and oral airflow rate. However, this may due to the influence of the manipulation of oral mask on the intensity level. Further investigation is needed before making a conclusion on this point.

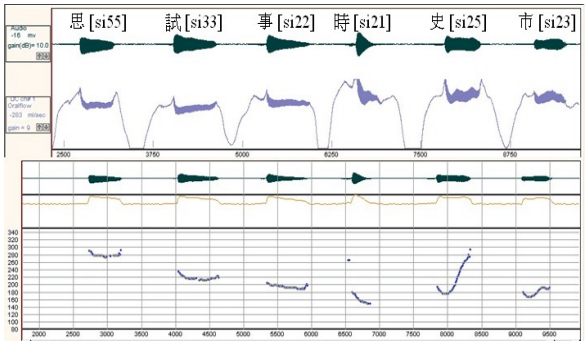

Figure 2: *The rate of oral airflow, intensity and F_0 contour for the six lexical tones associated with [si] in Cantonese.*

Upon inspection of the obtained data for intonation in this study, it shows that the findings at syllable level are also found at paragraph level. The variation in pattern of oral airflow rate and F_0 changes display a certain degree of similarity, despite the influence of tone and intonation interaction (Fig. 3). However, it should be noted that the pattern correlation between variation in the oral airflow rate and F_0 changes at paragraph level is not as good as that at syllable level. For syllable with a nasal ending associated with a high tone or rising tone and syllable under the influence of tonal co-articulation, the correlation between the variation in oral airflow rate and F_0 changes is poor. In Fig. 4, it shows that the F_0 contour of [kɔ33] at paragraph level is greatly influenced by the tones of the preceding and following syllables, which are [li55] and [jɐn25], respectively, whereas

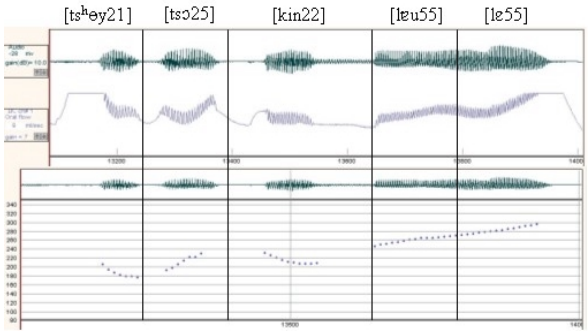

Figure 3: *The rate of oral airflow and F_0 contour for a clause within the paragraph [tsʰɵy21 tsɔ25 kin22 lɐu55 lɛ55] ' take the cloak off '*

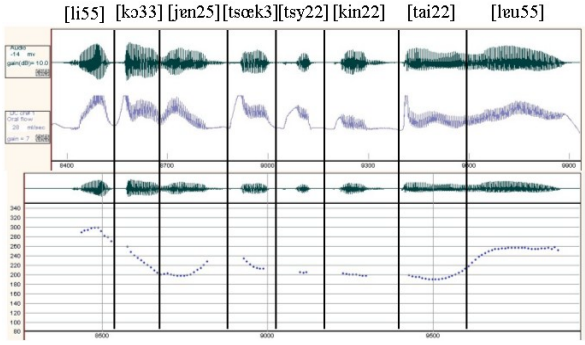

Figure 4: *The rate of oral airflow and F_0 contour for a clause within the paragraph [li55 kɔ33 jɐn25 tsœk3 tsy22 kin22 tai22 lɐu55] ' the traveller wore a warm cloak '*

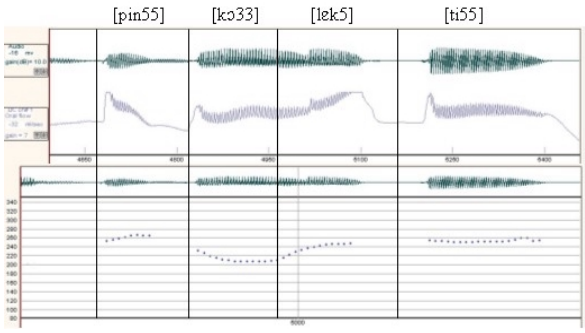

Figure 5: *The rate of oral airflow and F_0 contour for a clause within the paragraph [pin55 kɔ33 lɛk5 ti55] ' which is stronger '.*

the rate of oral airflow is independent of the tonal co-articulation. In Fig. 5, it can be seen that when the syllable with a nasal ending [pin] associated with a high level tone [55], there is a fall in the rate of oral airflow. It is different from the syllables in the same clause without nasal ending, [lɛk5] and [ti], that associated with a high level tone [55], of which the correlation between the rate of oral airflow and F_0 contour is quite good. As the pattern of the graphs for the three speakers are similar, all graphs that are shown in this

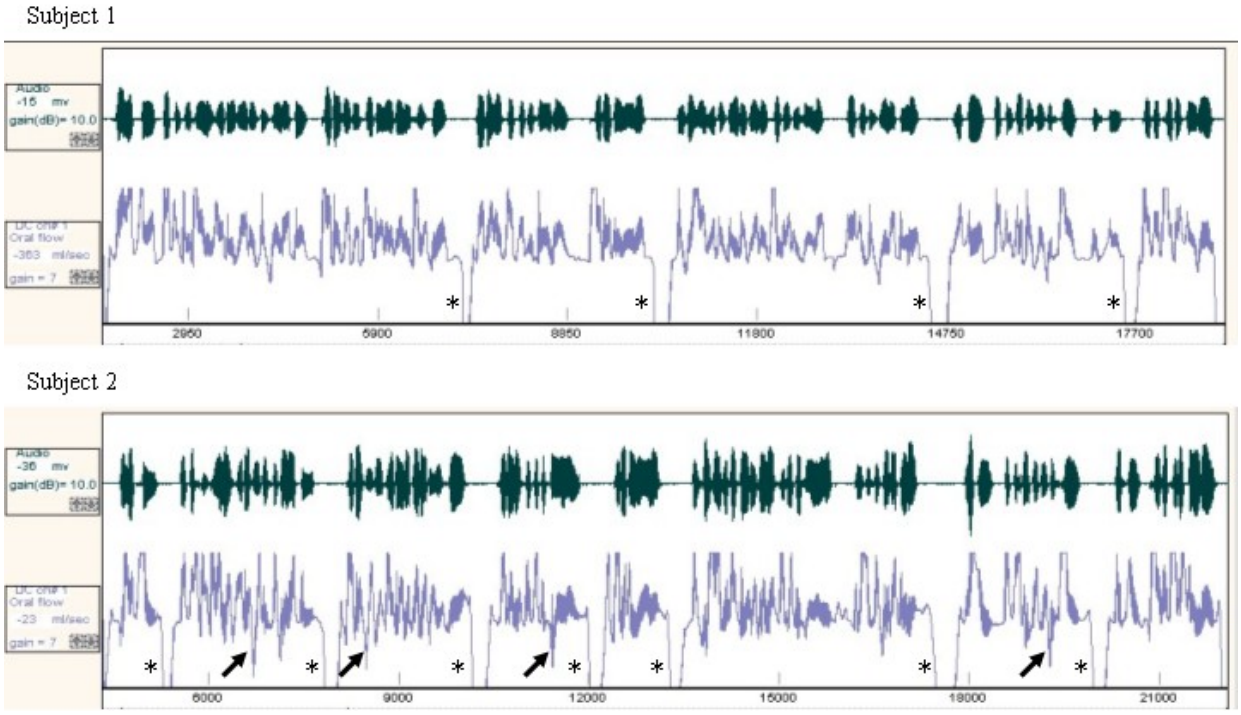

Figure 6: *The waveform and rate of oral airflow for the first half of the reading text produced by two of the three native speakers.*

paper for illustration are based on the data of one of the three speakers. Based on the logged data of F_0 and intensity, correlation coefficients were calculated between all possible pairs of these variables using the Pearson formula. It showed that in this study, the correlation between the F_0 and intensity is low.

3.2. Basic breath unit

Lieberman [9] proposed a theory of intonation claiming that intonation of a language was characterized by a basic breath-group or 'archetypal normal breath-group', which involved the archetypal pattern of articulatory activity. The archetypal pattern involved a state of least control on laryngeal muscles throughout expiration so that the changes in F_0 would be directly proportional to the subglottal pressure. Thus, the falling of subglottal air pressure would result in a rapid fall of F_0 at the end of a breath-group when inspiration took place. According to Liberman, the archetypal pattern produced a prosodic pattern that "is characteristic of ones that are used to delimit the boundaries of unemphatic, declarative sentences in normal speech" (p.27). It indicated that adult speakers in America English generally produce a breath-group on a single expiration.

Part of the aims of this study are to search for (if any) the domain of 'basic breath unit' in a to ne language and to examine if the 'basic breath unit' can be used to delimit the boundaries of declarative clauses at paragraph level. The 'basic breath unit' in this paper refers to the boundary within intonation at paragraph level where inspiration occurs, or in other words, where expiration ceases.

Fig. 6 shows the first half of the reading text produced by two of the three speakers. As can be seen, the trough indicates the boundary at paragraph level where inspiration occurs, as shown by an asterisk on the left. It shows that though the occurrences of inspiration mostly correspond to the clause and intonation boundaries with unfilled pauses, it is not the case that each clause and intonation boundary corresponds to a single expiration. In some cases, as indicated by the arrows, the inspiration corresponds to something smaller than a clause. It can also been seen that the boundary where inspiration occurs within paragraph varies from subject to subject. It shows that the boundary where inspiration occurs and the frequency of inspiration for each subject is not necessary the same. As shown in Fig. 6, the frequency of inspiration for Subject 1 is less than that for Subject 2. It seems that the domain of the 'basic breath unit' in Cantonese is not so clear and the 'basic breath unit' is speaker -specific.

3.3. Oral airflow rate and downtrend in F_0

A downtrend in F_0 has often been observed in running speech [1, 9]. Vance [15] indicated that the downtrend in F_0 does occur in Cantonese. As the F_0 is shown to be directly controlled by the changes in subglottal pressure, keeping everything else being constant, it is assumed that the downtrend in F_0 may due to a physiologically-based universal of gradual reduction of subglottal pressure throughout the course of an utterance. In this study, it shows that the downtrend in F_0 does not correspond to the variation in oral airflow rate. That is, the oral airflow rate for the first syllable within a sentence can be similar to that for the final syllable. It seems that though the subglottal pressure gradually decreases throughout the course of an utterance, some mechanisms are manipulated for keeping the oral airflow rate unchanged. As shown in Fig. 7, with the same citation tone, the oral airflow rate for the first syllable [pin55] is similar to that for the final syllable [lɛ55], even the level of F_0 contour

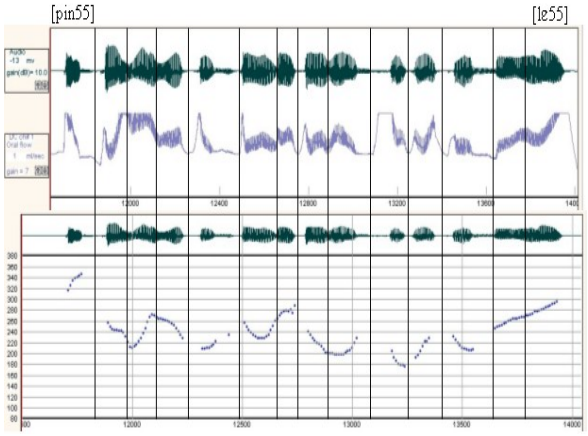

Figure 7: *The rate of oral airflow and F_0 contour for a clause within a paragraph [pin55 kɔ33 hɔ25 ji23 tsɪŋ25 tou33 li55 kɔ33 jɐn25 ts^{h}ɵy21 tsɔ25 kin22 lɐu55 lɛ55] ' the one who succeeded in making the traveller take his cloak off '.*

for the final syllable is lower than that for the first one under the influence of downtrend in F_0.

4. Discussion

The results indicate that the correlation in pattern between the oral airflow rate and F_0 changes is quite good at both syllable level and paragraph level. However, the changes in the level of F_0 contour do not correspond to the variation in oral airflow rate. It is plausible to assume that in spontaneous speech, the function of oral airflow is just to set the vocal cord vibrating and to facilitate the F_0 changes for rising or falling tone in some occasions, while the dominant factor in controlling the modulation of F_0 in speech is the laryngeal muscles [4, 10]. Owing to the same plausible reason, it is not difficult to understand why the pattern correlation between the oral airflow rate and F_0 contour is poor for syllables within paragraph that under the influence of tonal co-articulation. It seems to be reasonable that as the oral airflow rate does not contribute to the major F_0 changes, the influence of tonal co-articulation on the citation tone would overcome the influence of oral airflow rate. The poor correlation for syllable with nasal ending can also be explained by the physiological processes for the production of nasal sound. As the production of nasal sound involves lowering of the velum and the vowel in CVN syllable would be nasalized, airflow from the lung would escape through the velum to the nasal cavity and it results in a drop in the rate of airflow flowing out through the oral cavity. Results show that the 'basic breath unit' at paragraph level does not always correspond to a clause or sentence. The locations for the existence of 'basic breath unit' and the frequency of its occurrences are shown to be speaker-specific for the three speakers in this study. This indicates that the location for the occurrence of each 'basic breath unit' and the frequency of occurrences may correspond to the speaking style and basic physiological constraint of the respiratory system for a speaker.

5. Conclusions

This study describes a preliminary study on aerodynamic correlates of intonation in Hong Kong Cantonese at paragraph level. To conclude, in this study, the pattern correlation between F_0 contours and the rate of oral airflow is quite good at both syllable and paragraph level. However, the levels of F_0 contour for different citation tones are not controlled by the oral airflow rate. The function of oral airflow rate may be mainly for vocal cord vibrating and to facilitate certain F_0 changes. The correlations between F_0 and intensity, as well as oral airflow rate and intensity are shown to be poor. The results show that the 'basic breath unit' is not necessary correspond to every clause or sentence. The existence of 'basic breath unit' can correspond to units less than a clause. Moreover, the boundary where inspiration occurs and the frequency of inspiration throughout the course of an utterance vary from speaker to speaker. It may be due to the fact that the speaking style and basic physical need of inspiration for every speaker during speech are different. Results also show that the rate of oral airflow does not correspond to the downtrend in F_0, which may result form the decrease in subglottal pressure throughout the course of an utterance.

6. References

[1] Atkinson, J.E., 1978. Correlation analysis of the physiological factors controlling fundamental voice frequency. *Journal of the Acoustical Society of America* 63(1), 211-222.

[2] Baer, T., 1979. Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. *Journal of the Acoustical Society of America* 65(5), 1271-1275.

[3] Chao, Y.R., 1968. A grammar of spoken Chinese. Berkeley: University of California Press.

[4] Collier, R., 1975. Physiological correlates of intonation patterns. *Journal of the Acoustical Society of America* 58(1), 249-255.

[5] International Phonetic Association, 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

[6] Ladefoged, P., 1962. Subglottal activity during speech. *Proceedings of 4th International Congress of Phonetic Sciences*, 73-91.

[7] Ladefoged, P., 1963. Some physiological parameters in speech. *Language and Speech* 6, 109-119.

[8] Ladefoged, P., 1967. *Three areas of experimental phonetics*. Oxford: Oxford University Press.

[9] Lieberman, P., 1967. *Intonation, perception and language*. Cambridge: MIT Press.

[10] Ohala, J., 1970. Aspects of the control and production of speech. *UCLA Working Papers Phonetics* 15, 1-192.

[11] Shen, X-N. S., 1990. *The Prosody of Mandarin Chinese*. Berkeley: University of California Press.

[12] Strik, H.; Boves, L., 1992. Control of fundamental frequency, intensity and voice quality in speech. *Journal of Phonetics* 20, 15-25.

[13] Strik, H.; Boves, L., 1995. Downtrend in F_0 and P_{sb}. *Journal of Phonetics* 23, 203-220.

[14] Tseng, C.Y., 1990. *An acoustic phonetic study on tones in Mandarin Chinese*. Taipei: Institute of History and Philology, Academic Sinica.

[15] Vance, T.J., 1976. An experimental investigation of tone and intonation in Cantonese. *Phonetica* 33, 368-392.

[16] Yuen, I., 2002. Tonal invariance and downtrend in Cantonese. *Proceedings of Speech Prosody 2002*, France.