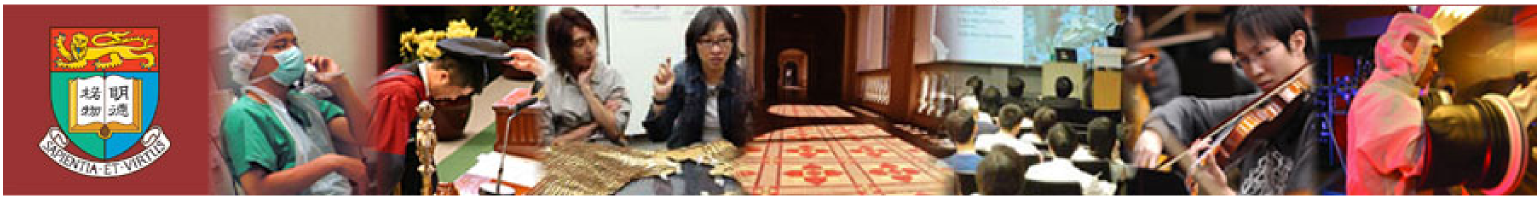


Title	**Interaction between lexical tone and labial movement in Cantonese bilabial plosive production**
Author(s)	**Tong, E; Ng, ML**
Citation	**The 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, 17-21 August 2011. In Proceedings of the ICPhS XVII, 2011, p. 2014-2017**
Issued Date	**2011**
URL	**http://hdl.handle.net/10722/136282**
Rights	**Creative Commons: Attribution 3.0 Hong Kong License**

INTERACTION BETWEEN LEXICAL TONE AND LABIAL MOVEMENT IN CANTONESE BILABIAL PLOSIVE PRODUCTION

Eric Tong & Manwa L. Ng

Speech Science Laboratory, Division of Speech and Hearing Sciences,
University of Hong Kong, Hong Kong
tong0923@hku.hk; manwa@hku.hk

ABSTRACT

The present study investigates the possible acoustic-articulatory interaction during production of Cantonese plosives at different lexical tones. Using electromagnetic articulography (EMA), movement of the jaw, and upper and lower lips was measured during production of word-initial bilabial plosives at six lexical tones. Preliminary data reveals a strong relationship between aspiration and lip opening velocity (LOV), as well as tone and range of labial separation. The present data indicate a dependency between the articulatory and phonatory systems of the human speech production mechanism.

Keywords: Cantonese, tones, lip, jaw, electromagnetic articulography

1. INTRODUCTION

The source-filter theory has been used to explain the mechanism of human speech sound production: the speech output is a product of the acoustic signal generated by vocal fold vibrations being modulated by the resonance characteristics of the supralaryngeal vocal tract [10]. The theory assumes the independence of the laryngeal (source) and supralaryngeal (filter) systems. However, validity of this assumption has been challenged in recent acoustic and kinematic studies.

Traditionally, kinematic information of speech articulators was obtained by using cineradiography or x-ray techniques, both of which have been proven to have potential hazards and could complicate the direct viewing and tracking the articulatory movements during continuous speech [9]. In recent years, electromagnetic articulography (EMA) has gained popularity and promoted studies testifying the independence relationship between the laryngeal and supralaryngeal systems [8]. EMA is a biologically safe, non-invasive technique that can be used to examine the real-time articulatory movement and obtain a large amount of kinematic data without interfering speech movements [4].

With the use of EMA technique, studies of the supralaryngeal and laryngeal interaction of a tone language have been done. For example, Hoole and Hu [7] studied articulatory gestures of the same vowels in Mandarin produced at different lexical tones using EMA. They found a significant effect of tone production on vowel articulation. Different lexical tones were associated with slightly different articulatory gestures during vowel production. This result provided insight about the relationship between the laryngeal system (represented by different lexical tones) and the supralaryngeal system (represented by varying articulatory gestures). In another study examining the relationship between lexical tone and articulation in Mandarin, Erickson, Iwata, Endo and Fujino [3] observed movement of tongue and jaw when the same syllable was produced at different lexical tones, and found significant difference in tongue and jaw movement for different tones. Based on these results, interactive patterns were found between the articulatory gestures and tone variations. It follows that there is a correlation between our laryngeal and supralaryngeal systems.

Findings from these studies appear to contradict the assumption of the source-filter theory. The present study attempted to extend these studies by examining the possible relationship between labial movements in bilabial plosive production at different Cantonese lexical tones.

During bilabial plosive production, a complete labial closure is formed and then released rapidly. According to Clark, Yallop and Fletcher [2], the entire process can be divided into three phases: (1) occlusion phase, (2) hold phase, and (3) release phase. Firstly, formation of a lip closure is completed during the occlusion phase. It is followed by the building up of a positive intra-oral pressure which is maintained behind the oral occlusion during the hold phase. As production of Cantonese bilabial plosives is egressive, an outward flow of airstream needs to be achieved. Finally, the labial occlusion is released rapidly by the forced opening of lips due to the high intra-oral pressure, i.e., release burst in release phase [2]. It should be noted that the release phase is followed by the onset of vocal fold vibration, which corresponds to phonation of the following vowel. Movement of the two lips is usually visible to the listener. This, in fact, serves as an important visual cue for the correct identification of bilabial plosives.

According to the assumption of source-filter theory, lip movement during production of bilabial plosive should in theory be independent of vocal fold phonation. Vocal fold behavior should not affect, or be affected by, the release of stop in the supralaryngeal vocal tract. In this study, lip movement of the same bilabial plosive produced at different Cantonese lexical tones was examined. The main research question of the present study was: *Does lip movement change with variations in tones in Cantonese, in a way similar to what has been observed in Mandarin?* Findings from the present study will help examine the validity of the laryngeal-supralaryngeal independence assumption of the source-filter theory.

2. METHOD

2.1. Participants

Three adult male native Cantonese speakers (aged from 19 to 23 years) participated in this study. A relatively small subject pool was used as EMA study is time-consuming and laborious. Yet, such sample size is comparable to many published studies in the literature. All participants were normal healthy speakers with no history of neurological disorders, oro-maxillo-facial surgery, and/or speech, language, or hearing problems. Contraindications of the instrument were avoided such as use of pace maker, claustrophobia and electromagnetic hypersensitivity. Only those who could speak naturally with the sensor coils attached were recruited.

2.2. Speech material

The speech material consisted of the monosyllabic Cantonese bilabial words /p^{h}a/ and /pa/ produced at the six Cantonese lexical tones: Tone 1 to Tone 6 (T1-T6). The six Cantonese tones represented by T1 to T6 are the high-level, high-rising, mid-level, low-falling, low-rising and low-level tones. During the experiment, Chinese characters corresponding to the monosyllables were printed on a sheet that was presented to the participants during the experiment. For the non-words /pa5/, /p^{h}a5/ and /p^{h}a6/, demonstrations were given to the participants and participants were allowed to practice until all the words and non-words could be produced correctly.

2.3. EMA preparation

The three-dimensional EMA system (AG500, Carstens Medizinelektronik) was used to measure the movement of upper and lower lips, and lower jaw during production of the Cantonese monosyllables. Warming up and calibration of the EMA system, and preparation and attachment of sensors were done strictly following the procedures described in the AG500 system manual before data collection.

During the EMA experiment, an alternating magnetic field of frequency from 7.5 kHz to 13.75 kHz was generated by the three pairs of transmitter coils in the EMA cube. They were used to induce signals in the small receiver coils which were attached on the lower jaw, upper and lower lips of participants upon movement. The magnitude of the induced signals in receiver coils was derived from the distance between the transmitter and transducer. Hence, movement of lips and jaw in z-y coordinate plane was obtained by measuring the position of receiver coils within a certain period of time. Displacement, velocity, acceleration and duration data were then calculated using the magnitude of the induced signals.

Receiver coils coated with latex were affixed to the lower jaw, upper and lower lips using biologically safe adhesive. Another three receiver coils were attached to the bridge of nose and the mastoid processes of the left and right temporal bones. All of them served as references to head movement.

2.4. Recordings

At the time of experiment, the participant was positioned in the center of the EMA cube with the sensor coils securely attached to the designated positions. They were instructed to produce 36 Cantonese monosyllabic words (2 syllables x 6 tones x 4 repetitions) at a comfortable loudness level and speech rate. To eliminate order effect, the order at which the stimuli were produced was randomized. Before the actual recording took place, the participants were given a five-minute practice period to accustom themselves with the recording environment and speaking with receiver coils attached. Acoustic data were recorded simultaneously for reference. The entire recording lasted for about 30 minutes for each participant.

2.5. Data analyses

Movement of the lower jaw, and upper and lower lips was measured during production of aspirated and unaspirated plosives /p^h/ and /p/ at different Cantonese tones. However, since changes in signal magnitude of receiver coils could also be contributed by head movement, correction for possible head movements relative to the EMA cube is needed. This was done to reposition movement data by taking references from the receiver coils at the bridge of nose and the two temporal bones. Kinematic data of the articulators included distance (mm), velocity (mm/s), and duration (ms) of the two lips. Mean values of each kinematic parameter were calculated by averaging the three productions of consonants.

Movement data were grouped by the two independent variables: aspiration (aspirated vs. unaspirated) and tones (T1-T6). Displacement and velocity of the two lips during production of /p^h/ and /p/ were compared.

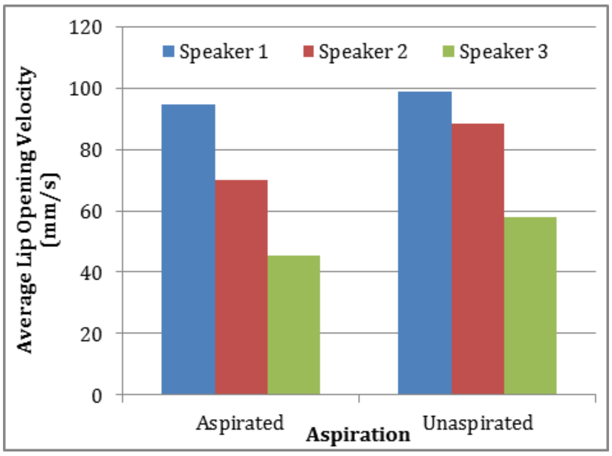
3. RESULTS & DISCUSSION

Lip opening velocity (LOV) indicates the average rate at which the lips open during the burst phase of plosive production. It is related to the time required for reaching the highest lip opening. LOV of each speech production was obtained by dividing the maximum separation of the two lips by the time elapsed.

3.1. Aspiration

As seen in Fig. 1, average LOV values of unaspirated plosives were found to be consistently higher than those of aspirated plosives. As Butcher [1] reported that aspirated plosive /p^h/ was consistently associated with a greater peak intraoral pressure than unaspirated plosive /p/ in various languages, faster lip movement should be resulted during the burst phase in aspirated plosives /p^h/. However, the present findings appear to contradict this suggestion; unaspirated plosives with lower intraoral pressure were found to have a faster lip opening, as indicated by the greater average LOV values. In a kinematic study of whispered bilabial plosives /p/ and /b/ of English, Higashikawa, Green, Moore, and Minifie [6] found that whispered /b/ was produced with significantly higher peak velocity in lip opening than whispered /p/. Suggested by Gracco [5], voiceless /p/ was recorded to have a slower lip opening velocity for formulating an appropriate voice onset time for better discrimination of laryngeal voicing or devoicing. To confirm this suggestion with Cantonese, further study of aerodynamics and physiology of lip opening during plosive production is required.

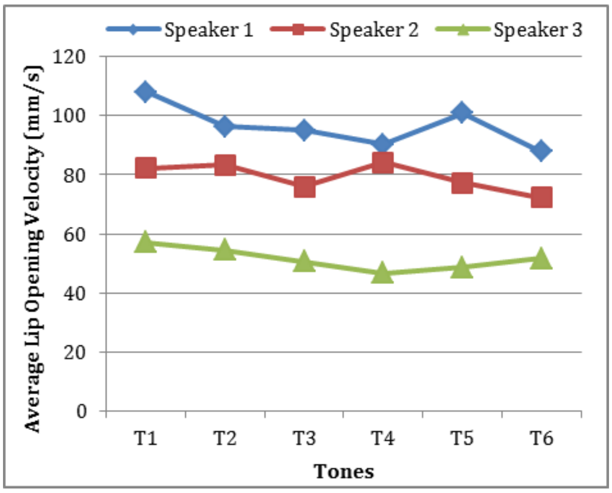
Figure 1: Mean lip opening velocity (LOV) of the aspirated and unaspirated productions of Cantonese bilabial words by three speakers.



3.2. Tones

To examine the effect of tones, comparison of LOV in six tone productions was done. Results are shown in Fig. 2.

Figure 2: Mean lip open velocity of six Cantonese tone productions of bilabial words by three speakers.



LOV of T1 (high level tone) was consistently higher than that of T6 (low level tone), with a generally decreasing pattern across the tones (see Figure 2). This implies that, for plosives produced at high tones, it took a shorter time to reach maximum lip separation than plosives produced at low tone. According to speech aerodynamics, high tones are produced a faster vocal fold vibration. In order vibrate the vocal folds more rapidly, both longitudinal tension of the vocal folds and subglottal pressure are found to be higher [11, 12]. With higher pressure beneath the vocal folds, more air will be released during the burst phase, and the lips should open to the maximum separation at a faster rate as the lips are now under a stronger opening force. Therefore, plosives produced at a higher tone are associated with a faster lip opening movement than those produced at a lower tone.

3.3. Limitations

3.3.1. Individual variations

As illustrated in Fig. 2, individual differences were found in LOV values and LOV pattern in tone productions among three speakers. Since the articulatory manner of the speakers was acquired and adapted, a considerable degree of variations in lip movement pattern would be expected. Therefore, these variations tend to limit the representativeness of the present results.

3.3.2. Non-word familiarity

Since there are three non-word stimuli (two aspirated and one unaspirated plosives), the participants had to practice for five minutes in order to gain familiarity with the non-words. However, accuracy of these word productions was not as high as those of real word and the articulatory may not be natural. As a result, it limited the accuracy of the LOV measurement.

To reduce the artifact associated with the use of non-words, it is suggested that longer practice time be given to the participants to practice the non-words until a natural production is achieved. This can increase the naturalness of the lip movement and hence a faithful LOV measurement.

4. CONCLUSION

The study yields preliminary results regarding the effect of aspiration and tones on articulatory kinematics during Cantonese bilabial plosive production. Kinematic data revealed that aspirated plosives are associated with slower lip opening, and high tone height in Cantonese show the faster labial separation. It is suggested that the slower rate of labial separation in aspirated plosives is related to difference in motor organization in aspirated plosives and unaspirated plosives, in order to maintain a distinguished difference in speech. For tonal height effect, the high subglottal pressure in high tones may contribute to the faster lip movement during the release phase of plosive production.

5. REFERENCES

[1] Butcher, A.R. 1992. Intraoral pressure as an independent parameter in oral stop contrasts. *Proceedings of the 4th Australian International Conference on Speech Science & Technology,* 286-291.

[2] Clark, J., Yallop, C., Fletcher, J. 2007. *An Introduction to Phonetics and Phonology.* Oxford: Blackwell.

[3] Erickson, D., Iwata, R., Endo, M., Fujino, A. 2004. Effect of tone height on jaw and tongue articulation in Mandarin Chinese. *Proc. Tonal Aspects of Languages* Beijing.

[4] Goozee, J., Murdoch, B.E., Theodoros, D.G., Stokes, P.D. 2000. Kinematic analysis of tongue movements in dysarthria following traumatic brain injury using electromagnetic articulography. *Brain. Injury* 14(2), 153-174.

[5] Gracco, V.L. 1994. Some organizational characteristicsof speech movement control. *Journal of Speech and Hearing Research* 37, 4-27

[6] Higashikawa, M., Green, J.R., Moore, C.A., Minifie, F.D. 2003. Lip kinematics for /p/ and /b/ production during whispered and voiced speech. *Folia Phoniatrica et Logopaedica* 55, 17- 27

[7] Hoole, P., Hu, F. 2004. Tone-vowel interaction in Standard Chinese. *Proc. Tonal Aspects of Languages,* Beijing.

[8] Hoole, P., Zierdt, A., Geng, C. 2003. Beyond 2D in articulatory data acquisition and analysis. *Proc. 15th ICPhS* Barcelona, 265-268.

[9] Perkell, J.S., Cohen, M.H., Svirsky M.A., Matthies, M.L., Garableta, I., Jackson, M.T.T. 1992. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements, *J. Acoust. Soc. Am.* 92(6), 3078-3096.

[10] Pickett. J.M. 2001. *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology.* Boston, MA: Allyn and Bacon.

[11] Plant, R.L., Younger, R.M. 2000. The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech. *Journal of Voice* 14(2), 170-177

[12] Titze, I.R. 1989. On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America* 85(2), 901-906.