Effect of Tone Height on Jaw and Tongue Articulation in Mandarin Chinese

Donna Erickson[1], *Ray Iwata*[2], *Mitsuaki Endo*[3], *& Akinori Fujino*[4]

[1]Gifu City Women's College, Gifu, 501-0192, Japan
[2]Kanazawa University, Kanazawa, Japan
[3]Aoyama Gakuin University, Tokyo, Japan
[4]NTT Communication Science Laboratories, NTT. Corp., Kyoto, Japan

[1]erickson@gifu-cwc.ac.jp

Abstract

This study investigates some of the articulatory changes that occur due to the tonal conditions of the syllable; specifically, jaw, tongue and formant frequency changes. Articulatory and acoustic data were collected using the electromagnetic articulatograph (EMA) at NTT Research Laboratories (Atsugi, Japan) from two native female speakers of Mandarin Chinese producing Tones 1 and 3 on monosyllabic words containing the vowel /a/ (ba, ma, pa). Measurements of both acoustics (F0, F1) and pellet positions (tongue tip, tongue dorsum, and jaw) were made at the center of the syllable, i.e., at the time when the jaw was lowest.

T-tests show that supralaryngeal articulation is significantly different for the two tones. For the low Tone 3 compared to high Tone 1, for both speakers, the jaw and tongue are significantly more retracted and F1 is significantly higher.

This study provides useful data for exploring the interactive control between laryngeal mechanisms and articulation of vocal tract shape, which a speaker of a tone language must access to produce tones. It also provides interesting data applicable to current issues in phonology and phonetics about the interaction between prosody and articulation.

1. Introduction

Currently there is much interest among phonologists and phoneticians about the effect of prosody on articulation. A conventional way of treating tones and vowels in phonology is as phonemic units. Within this viewpoint, it is often assumed that the vowel's formant frequencies will be constant across tonal contrasts. However, work with other languages, such as English, for instance, has shown that formant frequencies change due to stress conditions of the syllable. Specifically, Erickson (2002) suggested that stress-related changes in formant frequencies were a consequence of prosodically-induced changes in articulation (jaw and tongue movement patterns). This viewpoint is in keeping with that proposed by Fujimura's C/D Model of articulatory implementation in which syllable characteristics are strongly affected by the prosodic input to the spoken utterance (e.g., Fujimura, 2000). As concerns a tone language, such as Mandarin Chinese, what are some articulatory and formant frequency changes that occur due to the tonal conditions of the syllable?

An essential acoustic correlate of Mandarin tones is their F0 pattern (e.g., Howie, 1974). For instance, Tone 1 is characterized by a high F0, Tone 3, by a low F0. Tone 3 also may be accompanied by laryngealization toward the middle of the syllable (Hirayama, 1975; Belotel-Grenié& Grenié, 1997). Articulatory studies have shown that the F0 underlying tonal contrasts in Chinese are primarily due to changes in laryngeal muscle tension (e.g., Hallé et al., 1990). Specifically, high F0 is the result of tensing the vocal folds, primarily by using the cricothyroid muscle. Low F0 (below the speaker's mid range) may involve extra laryngeal strap muscles, i.e., sternohyoid, sternothyroid and thyrohyoid (see Hallé, 1994; Sagart et al., 1986; Iwata et al. for Suzhou tones, 1991; and Erickson for Thai tones, 1976, 1991).

The changes in F0 underlying tonal contrasts are recognized as primarily due to changes in laryngeal muscle tension and basically independent from supralaryngeal articulation. However, recent studies show there also may be supralaryngeal articulatory changes during the production of tones in Mandarin Chinese. Specifically, Torng (2000) has shown an interaction between supralaryngeal articulation and laryngeal control, which may account for intrinsic F0 differences of high and low vowels.

In this study, we investigate the following three questions: (1) What is the effect of tone height on jaw and tongue articulation? (2) What is the effect on formant frequencies? (3) What is the relationship of F0 of tones to jaw and tongue articulation of tones?

2. Methods

Articulatory and acoustic data were collected using the electromagnetic articulatograph (EMA) at NTT Research Laboratories (Atsugi, Japan) from two native female speakers of Mandarin Chinese producing tones on monosyllabic and disyllabic words containing the vowel /a/ as in /ba/, /ma/, /pa/. This paper focuses only on Tones 1 and 3, as representative of the extremes of F0 change, and only monosyllabic words. For the EMA data analysis, we examined the x-y tracings recorded from the movement of the receiver coils attached to (1) the surface of the lower incisor (as representative of the jaw position) and (2) the four receiver coils (T1, T2, T3, T4) attached along the longitudinal sulcus of the speaker's tongue (see Figure 1).

Fig. 1. *Placement of coils on jaw and tongue (T1, T2, T3, T4).*

Measurements of both acoustics (F0 and F1) and coil positions (T1, T2, T3 and jaw) were made during the syllable at the time when the jaw was at its lowest position. T4 was not measured because this coil was dislodged during recording for one of the speakers. Figure 2 shows a sample display of the acoustic and articulatory data. The vertical line running perpendicular through the syllable indicates the point in the syllable where measurements were made. For Tones 1 and 3, the maximum jaw opening was generally at the center of the syllable.

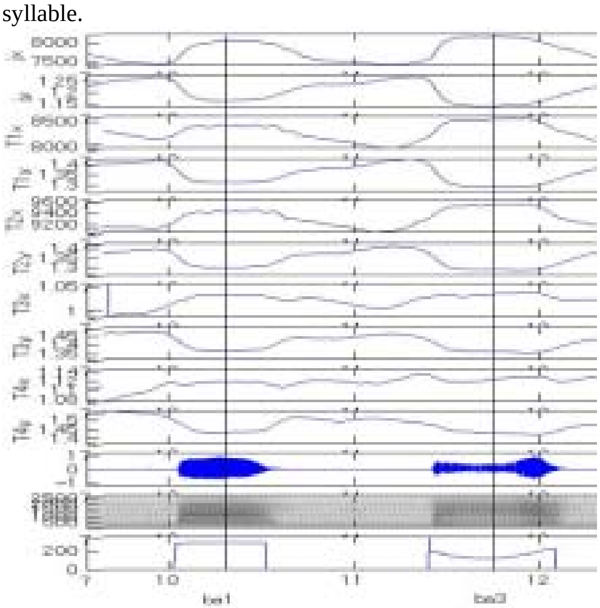

Fig. 2. *Sample articulatory and acoustic data for /ba1/ and /ba3/. In the top panels, jx and jy indicate jaw-x and jaw-y positions; the next 8 panels indicate the tongue x-y positions. Increased x-values indicate greater retraction; decreased y-values, greater lowering. Spectrograms and F0 are shown in the bottom two panels.*

3. Results

3.1 Effect of tone height on F0

As shown in Table 1, the F0 mean, as measured at the time of maximum jaw opening, was significantly higher for Tone 1 than Tone 3. This is as expected, since Tone 1 is a high tone, and Tone 3, a low tone.

Table 1. Mean F0 of Tones 1 and 3 for 2 speakers

Speaker	Tone 1	Tone 3
1	280.5 Hz (N=13)	120.6 Hz (N=14)
2	271.8 Hz (N=21)	182.3 Hz (N=18)

3.2 Effect of tone height on tongue dorsum articulation

In addition, the tongue is more retracted for Tone 3 than Tone 1. Mean values of T1, T2, and T3 are shown for Tones 1 and 3 for each of the speakers in Table 2 below. A double asterisk indicates significance of a one-sample t-test testing against the null hypothesis of mean difference = 0 with alpha= 0.01; a single asterisks, with alpha = 0.05. For both speakers T2 is significantly more retracted for Tone 3 than Tone 1 ($p<.01$) and for speaker 1, the tongue tip (T1) is significantly more retracted ($p<.05$). For Speaker 2, the tongue dorsum (T3) is significantly more retracted ($p<.01$).

Table 2. Mean tongue x-positions (mm) for Tones 1 and 3

Speaker 1			**Speaker 2**		
	Tone 1	Tone 3		Tone 1	Tone 3
T1x*	8.51	8.59	T1x	8.72	8.93
T2x**	9.44	9.54	T2x**	9.83	10.06
T3x	10.32	10.36	T3x**	10.86	11.11

The x-y positions of the tongue for Tones 1 and 3 are shown in Figs. 3, 4, and 5. below.

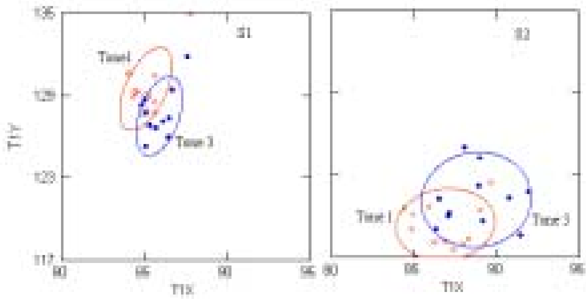

Fig. 3. *X-y tongue positions (T1) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel). The x-y axis values (mm) are displayed so that the leftmost bottom corner of the figures represents forward and downward position of the tongue and jaw, as if the speaker were facing to the leftside of the page.*

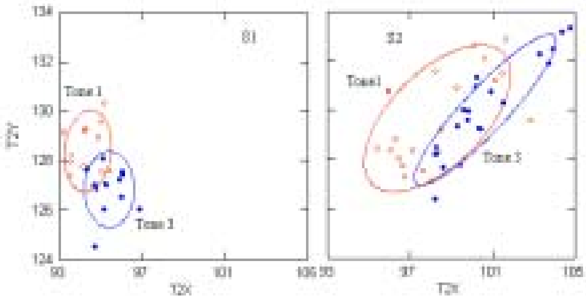

Fig. 4. *X-y tongue positions (T2) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel). Axes are the same as for Fig. 3.*

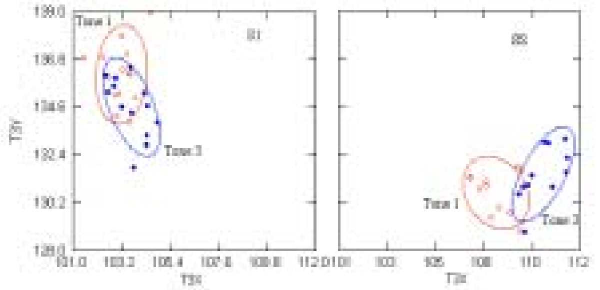

Fig. 5. *X-y tongue positions (T3) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel). Axes are the same as for Fig. 3.*

For both speakers, we see the tongue is more retracted for Tone 3 than for Tone 1. For Speaker 1 especially, the tongue is also lower for Tone 3 than Tone 1, as shown in Table 3 below.

Table 3. Mean tongue y-positions (mm) for Tones 1 and 3

Speaker 1 | **Speaker 2**

	Tone 1	Tone 3		Tone 1	Tone 3
T1y**	129.5	127.5	T1y*	119.4	121.2
T2y**	128.4	126.8	T2y	128.1	128.6
T3y**	136.1	134.6	T3y	130.6	131.5

3.3 Effect of tone height on jaw articulation

In addition to tongue retraction for Tone 3, the jaw is also more retracted. Mean jaw-x values for Tones 1 and 3 for each speaker are shown in Table 4 below. Speaker 2 especially shows jaw retraction for Tone 3. A t-test shows that the difference between Tone 1 and Tone 3 is slightly more significant for S2 ($p<.01$) than for S1 ($p<.05$).

Table 4. Mean jaw-x positions (mm) for Tones 1 and 3

	Tone 1 (jaw-x)	Tone 3 (jaw-x)
Speaker 1*	8.04 mm	8.12 mm
Speaker 2**	7.02 mm	7.10 mm

As for vertical position of the jaw, only for Speaker 1 is the jaw lower for Tone 3 than Tone 1. This difference is significant ($p<.01$), as shown in Table 5 below.

Table 5. Mean jaw-y positions (mm) for Tones 1 and 3

	Tone 1 (jaw-y)	Tone 3 (jaw-y)
Speaker 1 **	11.5 mm	11.31 mm
Speaker 2	11.54 mm	11.59 mm

3.4 Effect of tone height on F1

F1 is significantly higher for Tone 3 than Tone 1 for both speakers ($p<.01$) as shown in Table 6 below.

Table 6. Mean F1 values (Hz) for Tones 1 and 3

	Tone 1	Tone 3
Speaker 1 **	1157 Hz	1212 Hz
Speaker 2 **	1113 Hz	1162 Hz

3.5 Interspeaker differences with respect to lowering and backing of articulators

Although both speakers show increased F1 for the low Tone 3 compared with the high Tone 1, it is only Speaker 1 who lowers the jaw significantly more for Tone 3, as can be seen in Fig. 6 below. This speaker shows a positive correlation between amount of jaw opening and F1. For this speaker, the lower the jaw, the higher was the F1.

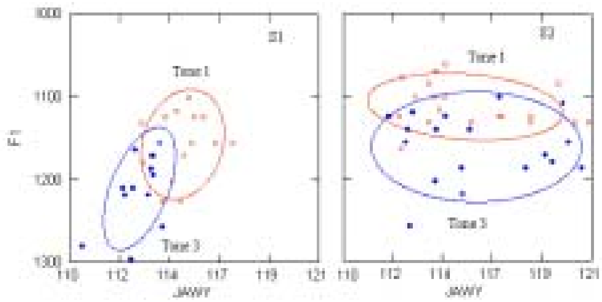

Fig. 6. *F1 (Hz) and jaw-y (mm) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel). The F1 axis is reversed in order to show the vowel space; the articulatory axis is the same as in Fig. 3.*

Fig. 7 shows that for Speaker 1 in producing Tone 3 the change in vertical position of the jaw is more pronounced than that of the horizontal position, whereas for Speaker 2, it is primarily the horizontal position that changes. (See also Tables 4 and 5.)

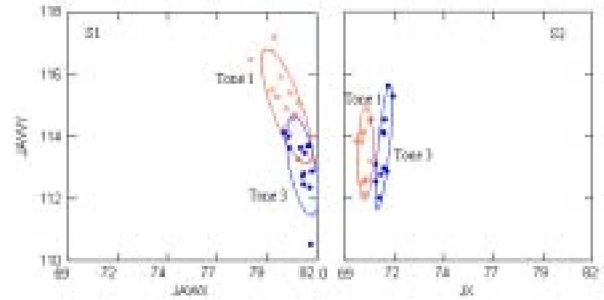

Fig. 7. *X-y jaw positions (T3) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel). Axes are the same as in Fig. 3.*

3.6 Interspeaker differences with respect to jaw-F0 relationship

Fig. 8 below shows a comparison of the productions of Speaker 1 *vs.* Speaker 2 of Tone 3 *vs.* Tone 1 in terms of jaw-y and F0 values. We see that for Speaker 1 compared to Speaker 2, the jaw position for low Tone 3 is lower than for Tone 1; moreover, the difference in mean F0 between the low Tone 3 and high Tone 1 is much greater (160 Hz). Compared to Speaker 2, the high tone for Speaker 1 is higher, and the low tone is lower. For Speaker 2, the jaw position does not change, and the difference in mean F0 is not so large (90 Hz).

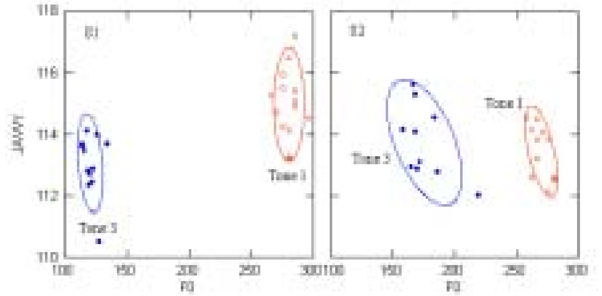

Fig. 8. *Jaw-y (mm) and F0 (Hz) for Tone 1 and Tone 3 for Speaker 1 (left panel) and Speaker 2 (right panel).*

4. Discussion

To produce the low Tone 3 in contrast to the high Tone 1, the tongue and jaw are more retracted, and F1 is raised. Retraction of tongue and jaw in the presence of low F0 has been reported previously by Honda et al. (1999). For low back vowels, it is compatible with articulatory modeling work by Dang and Honda (1997) that the acoustic consequence of changing the vocal tract space by retracting the jaw and tongue, (i.e., by increasing the ratio of the frontal and back cavities) is a raised F1. In addition, the retraction of the articulators may contribute to the characteristic laryngealization associated with Tone 3. The tongue-jaw mechanism is also involved in the intrinsic F0 of low vowels (e.g., Zawadzki and Gilbert, 1989), as also reported by Torng (2000) for Mandarin for high and low vowels.

Interesting interspeaker differences in terms of effects of tone height on articulation are seen in this data. For Speaker 1 to produce Tone 3 *vs.* Tone 1, both the jaw and F0 are lower than that seen for Speaker 2. Speaker 1 tends to change vertical jaw position whereas Speaker 2, changes horizontal tongue dorsum position.

5. Conclusion

For both speakers, we see changes in articulation due to the tonal conditions of the syllable. For the low Tone 3, the tongue and jaw are more retracted and F1 raised. This may be explained by the tongue-jaw-larynx linkage, such that during low F0, the jaw and tongue are retracted. It also may be related to the laryngealization characteristic of Tone 3. This needs to be investigated further.

In addition, there are interspeaker differences in jaw articulation. The speaker who showed the lowest F0 for Tone 3 also tended to lower the jaw considerably more than the other speaker for this tone, who primarily retracted the jaw (and tongue).

Our current project involves investigating other tones and other vowels--the effect of tone height on not only low vowels (as in this study), but also high vowels. We are also examining the effect of lexical and phrasal stress, as well as syllable position, on articulation. Examination of a large number of speakers is very interesting given the interspeaker variations found in this study.

The fact that both vocalic and tonal specifications interact to yield the phonetic constellation of the syllable suggests that it may be useful to think of the syllable, rather than phonemes, as the domain of linguistic information. Our findings about the effect of tone height on supralaryngeal articulation suggest an interaction between phonetics and phonology that needs further exploration.

6. Acknowledgements

We thank NTT Communication Science Laboratories, Atsugi, Japan for use of the EMA facilities and Jianwu Dang, ATR Kyoto Japan for the MATLAB analysis program. Research was made possible in part by a grant to the second author from the Japanese Ministry of Education. This paper is an expanded version of a paper presented at the 11th Annual Conference of the International Association of Chinese Linguistics, 2002.

7. References

Belotel-Grenié, A. & Grenié, M. (1997). Types de phonation et tons en chinois standard. *Cahiers de Linguistique-Asie Orientale* 26 (2), 249-279.

Dang, J., and Honda, K. (1997). Acoustic characteristics of the piriform fossa in models and humans, *J. Acoust. Soc. Am*. 101, 456-465.

Erickson, D. (1993). Laryngeal muscle activity in connection with Thai tones. *Annual Bulletin Research Institute Logopedics and Phoniatrics, Univ. of Tokyo*, 27, 135-149.

Erickson, D., 2002. Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59, 134-149.

Howie, J., (1974), On the Domain of Tone in Mandarin. *Phonetica* 30, 129-148.

Fujimura, O. (2000). The C/D model and prosodic control of articulatory behavior. *Phonetica* 57, 128-38.

Hallé, P. A. (1994). Evidence for tone-specific activity of the sternohyoid muscle in modern standard Chinese. *Language and Speech*, 37 (2), 103-29.

Hallé, P.A., Niimi, S., Imaizumi, S., & Hirose, H. (1990). Modern Chinese 4 tones: EMG and acoustic patterns revisited. *Ann.Bull.RILP*, 24, 41-58.

Hirayama, H. (1975). Pekin-go no seicho taikei—tokuni dai-3 sei no kaishaku o megutte. *Gengonokagaku*, 5, 85-96.

Honda, K., Hirai, H., Masaki, S., and Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in F0 control. Language and Speech, 42, 401-411.

Iwata, R., Hirose, H., Niimi, S., and Horiguchi, S. (1991) Physiological properties of "Breathy" phonation in a Chinese dialect - A Fiberoptic and Electromyographic study on Suzhou dialect – *Proceedings of XIIth Int'l. Conf. on Phonetic Sciences*, Vol.3, 162-165.（Chinese translation appeared in F.Shi ed. *Hanyu Ynajiu Zai Haiwai* (1995),.48-58.）

Kaburagi, T., and Honda, M. (1997). Calibration methods of voltage-to-distance function for an electromagnetic articulometer (EMA) system. *J. Acoust.Soc. Am*., 101, 2391-2394.

Kaburagi, T., and Honda, M. (2002). Electromagnetic articulograph based on a nonparametric representation of the magnetic field. *J. Acoust.Soc. Am*., 111, 1414-1421.

Sagart, L., Halle, P., de Boysson-Bardies, B., & Arabia-Guidet, C. (1986). Tone production in Modern Standard Chinese: an electromyographic investigation. *Cahiers de Linguistique Asie Oreintale*, 15 (2), 205-221.

Torng, P-C (2000). *Supralaryngeal articulator movements and laryngeal control in Mandarin Chinese tonal production.* Doctoral Thesis, University of Illinois, Urbana-Champaign.

Zawadzki, P. A., and Gilbert, H. R. (1989). Vowel fundamental frequency and articulator position. *J. Phonetics*, 17, 159-166.