AN ELECTROMYOGRAPHIC INVESTIGATION OF LARYNGEAL MUSCLE ACTIVITY IN MODERN STANDARD CHINESE TONES

Laurent SAGART — Centre de Rech. Ling. sur l'Asie Orientale, CNRS, EHESS. Paris, FRANCE

Pierre HALLE — Labo. de Psychologie expérimentale. CNRS, EHESS. Paris, FRANCE

Bénédicte de BOYSSON-BARDIES — Labo. de Psychologie expérimentale CNRS, EHESS. Paris, FRANCE

ABSTRACT

The activity patterns of the cricothyroid and sternohyoid muscles in Modern Standard Chinese tones were electromyographically investigated in two subjects. Average activity profiles by tone and by segmental syllable indicate cricothyroid activity is well correlated to Fo with a latency time of 80-100 ms. The sternohyoid participates both in Fo lowering and segmental articulation, with strong activity peaks preceding consonant release when the following vowel is back and/or low.

INTRODUCTION

The present study is part of an ongoing electromyographic (EMG) investigation <14> of modern standard Chinese (MSC) tones. In terms of the pitch contours which characterize them, the four lexical tones of MSC may be roughly described as: tone 1 (T1): high level, tone 2 (T2): mid to high rising, tone 3 (T3): mid-low to low falling, tone 4 (T4): high to low falling. for detailed acoustic descriptions see <7,10,11>. The activity patterns of two laryngeal muscles: the Cricothyroid (CT), the main regulator of vocal fold tension, and the Sternohyoid (SH), an extrinsic laryngeal muscle shown by various EMG studies <13,5> to be active in Fo falls and low Fo, but also believed to be involved in segmental articulation, were investigated. The subjects were two female students in linguistics in their late twenties, both native speakers of MSC. Subject CYC came from Taipeh and subject FJQ from Beijing. The corpuses consisted of MSC syllables in all tones placed in a carrier sentence ([wo3 niæn4 X tsɨ4] "I read the character X") to avoid contamination by non-speech muscular activity. The target syllables were meaningful MSC words belonging to minimal series having the same segmentals, with lexical items in the four tones. The same minimal series were used with both subjects, and additional material was also used for subject FJQ.

METHOD

Two thin hooked platinum wire electrodes were percutaneously inserted into both subjects' CT and SH muscles according to a technique described in <3>. The subjects were then made to perform various manoeuvers (opening jaw against opposing force, swallowing, holding breath) to check on the insertion of the electrodes in the desired muscles. As the subjects read the corpuses aloud in a sound-treated room, the EMG signal from the electrodes and the acoustic signal were simultaneously recorded by means of a 7-track AMPEX recorder. Due to displacement of some electrodes after the checking manoeuvers, the signal from subject CYC's SH muscle proved unpossible to interpret. As a result, only signals from subject FJQ's CT and SH and subject CYC's CT could be analyzed.

The audio and EMG signals were digitized at 8 KHz (after low-pass filtering at 3.5 KHz and 6 dB/ octave analog preemphasis for the audio signal) and stored on disk in a Solar 16-40 computer. Fo was extracted by means of a cepstral method with framelength set to 312 ms. and frame period set to 10 ms. The EMG signal was undersampled to 1 KHz, and the absolute values were then integrated over a 75 ms. Hamming window sliding by 4 ms. steps. Programs were designed for the purpose of displaying the audio and EMG signals together with the Fo and integrated EMG curves. The tone-carrying part, consisting of the main vowel and any segment following it in the same syllable <8> was visually identified on the audio tracing. Based on a technique created by Kratochvíl <9> for obtaining average Fo and Ao profiles for tones, Fo and integrated EMG were measured by hand in twelve regularly spaced points of time (numbered -2 to 9) for each target-syllable, points 1 and 6 corresponding to the onset and endpoint respectively of the tone-carrying part as determined on the audio tracing, with intervals of 20% of the duration of the tone-carrying part between any two adjacent points. These 24 measurements summarized the evolution of the EMG and Fo curves for each tone-carrying part and the margins on both sides of it. Mean values and standard deviations were calculated for each of these 24 points to obtain average Fo and EMG profiles by tone and by segmental syllable.

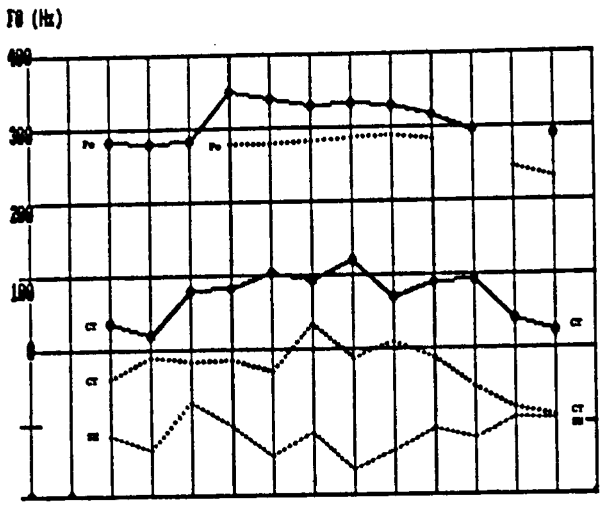

Fig. 1.a: tone 1.

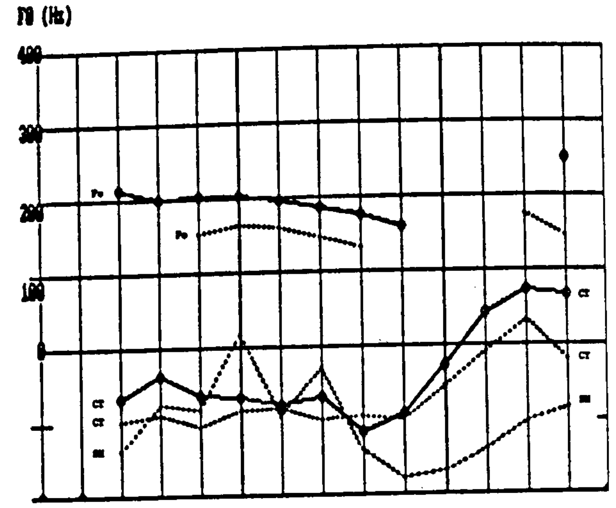

Fig. 1.c: tone 3.

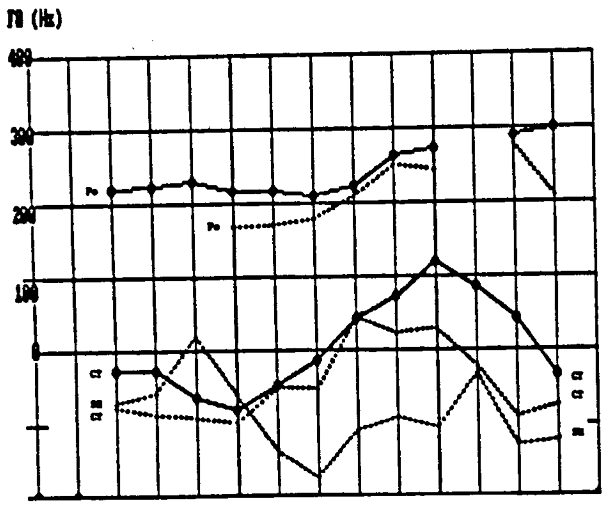

Fig. 1.b: tone 2.

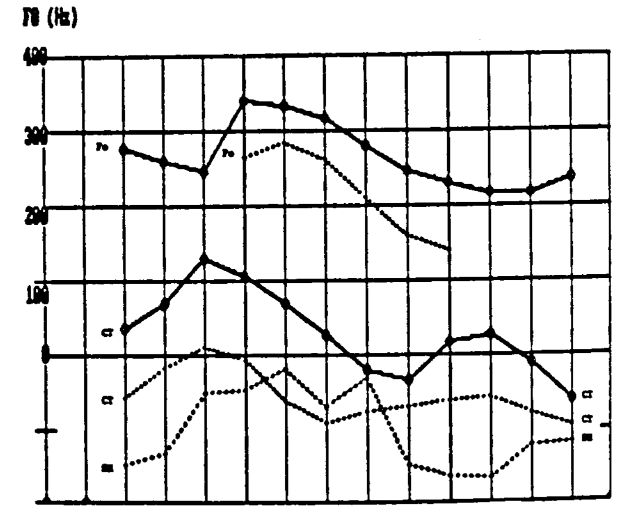

Fig. 1.d: tone 4.

Figs. 1.a-d: average Fo, cricothyroid and sternohyoid activity profiles by tone. Each curve averages Fo or muscle activity in five test syllables with segmentals [ma, ?ai, u, tu, thu]. Solid line: subject CYC. Dotted line: subject FJQ. Thick vertical lines indicate onset (left) and endpoint (right) of tone-carrying part.

RESULTS AND DISCUSSION.

Figs. 1.a-d show Fo and muscle activity profiles for each tone. With speaker FJQ, the CT and SH profiles were obtained from the same tokens.

The cricothyroid: similar evolutions of the CT are observed with both subjects: Fo rises (second part of T2, and any rises preceding the onsets of T1 and T4) are preceded by increases in CT activity. High CT activity also accompanies T1, a tone with high Fo throughout. In contrast, portions of tones characterized by Fo falls (T3, T4) are preceded by decreases in CT activity.

Latency time, the interval between muscle activity and Fo response, is best estimated by cross-correlation methods but a rough estimate can be arrived at by measuring the interval between remarkable points on the integrated EMG curve and the corresponding points on the Fo curve. For both subjects, latency times thus measured range between 50 and 160 ms, with values most frequently situated in the 80-100 ms range. This is in agreement with the finding in <4> of a mean latency time of 94 ms. for this muscle. In Figs.1.a-d, these values correspond to between one and two times the interval between two adjacent points. Accordingly, patterns of CT acti-

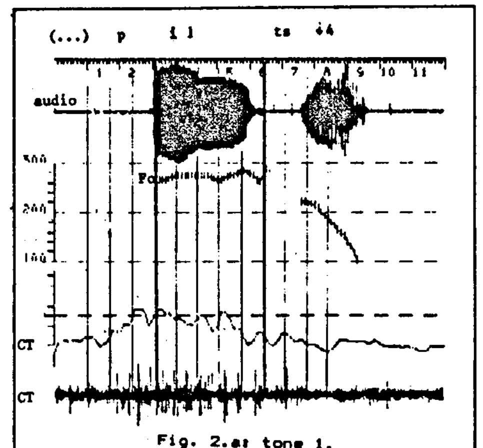

Fig. 2.a: tone 1.

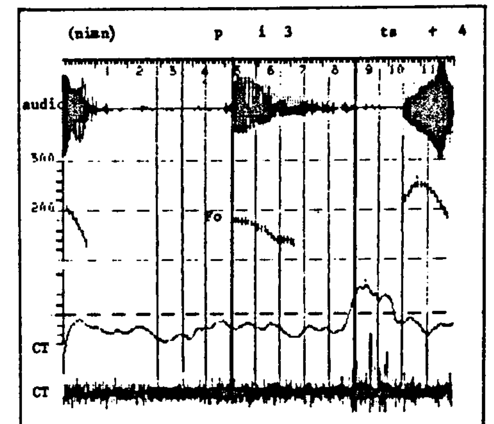

Fig. 2.c: tone 3.

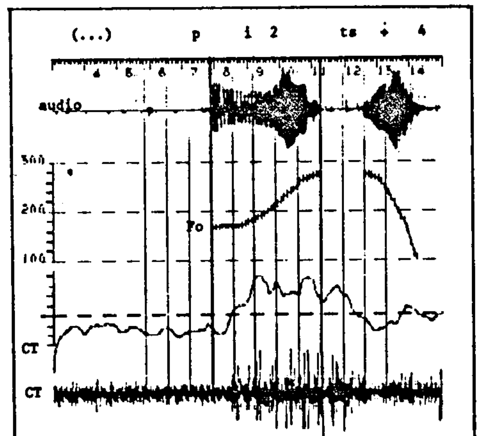

Fig. 2.b: tone 2.

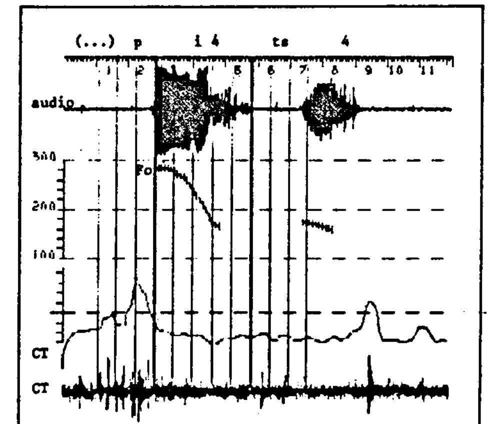

Fig. 2.d: tone 4.

Figs 2. a-d: Fo and cricothyroid activity in syllable [pi] in the 4 tones (subject FJQ). Test syllables are inserted in carrier sentence: [wo3 niæn4 X tsɨ4] "I read the character X". From top to bottom: audio signal, Fo, integrated EMG curve and raw EMG signal.

vity occurring on points -1 to 1 are relevant to the control of Fo at the beginning of the tone-carrying part of the test syllable, but patterns occurring after point 4 cannot relate to the test syllable. Note the increase in CT activity in the vicinity of point 6 in figs 1. a-d. This, we believe, relates to the production of T4 in the following syllable, "tsɨ4".

Average profiles by segmental syllable indicate no clear effect of segmentals on the level of CT activity (although syllable structure may affect the location of CT peaks in relation to segmental events). In particular we do not observe the correlation between vowel timbre and peak level of CT which Auteserre & al. <2> suggested might account for the intrinsic frequency of vowels.

This overall pattern is stable across utterances, and also across repetitions of the same utterance. It characterizes not only the test syllables but also the carrier sentences. It corresponds well to the activity pattern of the CT as described for other tonal or non-tonal languages: Swedish <6>, Dutch <4>, Thai <5>, French <2> etc. Typical examples of CT activity in the 4 tones are shown in figs. 2.a-d.

The sternohyoid. In spite of very high standard deviation values, the average SH profiles by tone in figs. 1.a-d show some correlation between low or falling Fo and increased SH activity, in particular in T3 and T4. Conversely, low SH activity accompanies T1, the high level tone, and T2, the rising tone. However, the Fo shoulder at the end of T2 is often preceded by an SH peak.

All profiles regardless of tone also display an increase between points -2 and 0, corresponding to sharp activity peaks shortly before vowel onset in individual test syllables. This suggests that part of the activity of the SH is unrelated to pitch control. Average profiles by segmental syllable (Fig.3) indicate the level of activity around vowel onset depends on the nature of segmental material, vowel timbre in particular: highest levels occur with [a], lowest with [i], intermediate levels with [u]. These observations of the activity pattern of the SH are consistent with published accounts of its role in speech. Regarding Fo control, Ohala <12> claimed the strap muscles, among which the SH, lower Fo indirectly by lowering the larynx, which in turn reduces the vertical, not the antero-posterior, tension in the vocal folds. Regarding segmental articulation, Ohala and Hirose <13> claimed that the SH also participates in tongue-lowering, tongue-backing and jaw-opening gestures by fixing or lowering the hyoid bone when muscles linking the hyoid bone and structures above it are also contracting (as the anterior belly of the digastric in jaw-opening gestures). Fo lowering will occur only if the hyoid bone and the larynx are free to move downward, that is, if the hyoid bone is not simultaneously pulled upward by muscles above it.

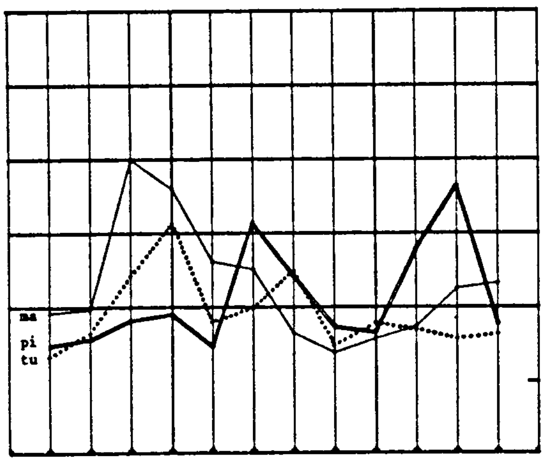

Fig.3: Average sternohyoid activity profiles by segmental syllable (subject FJQ). Each curve averages SH activity in four test syllables with T1, T2, T3 and T4. Simple line: syllable [ma]. Dotted line: syllable [tu]. Double line: syllable [pi].

CONCLUSION

Involvement of the SH in both Fo control and segmental articulation is not the only source of variability in our data: SH activity patterns can differ widely in repetitions of the same utterance. While instability of muscle activity patterns is a normal result in EMG studies, the stability of CT patterns is more remarkable. Two reasons may be invoked to account for it: (a) the CT specializes in Fo control and is not simultaneously involved in other tasks, and (b) although other muscles (among which the SH) play a role in Fo control, none is so efficient in regulating vocal cord tension.

ACKNOWLEDGEMENTS

The EMG recordings were made at INSERM (U3), Hôpital de la Salpêtrière, Paris, under the supervision of Dr. Chevrie-Muller. The electrodes were inserted by Dr. Lacau Saint-Guily and the signal-collecting apparatus was installed by Mr. B. Maton. The authors gratefully acknowledge the assistance received.

REFERENCES

<1> Atkinson, J.E. (1973) Physiological factors controlling Fo: results of a correlation analysis. Abstract in JASA 54, 1.319 (A).

<2> Auteserre, D., A. Di Cristo & D. Hirst (1986) Approche physiologique des intonations de base du francais: cricothyroidien et fréquence fondamentale. Proceedings of 15èmes Journées d'étude de la parole (Aix-en-Provence, 1986), pp.37-41.

<3> Basmadjian, J.V. and G.A. Stecko (1962) A new bipolar indwelling electrode for electromyography. Journal of applied physiology, 17, p. 859.

<4> Collier , R. (1975) Physiological correlates of intonation patterns. JASA 58,1:249-255.

<5> Erickson, D. (1976) A physiological analysis of the tones of Thai. U. of Connecticut PhD.

<6> Gårding, E., O. Fujimura & H. Hirose (1970) Laryngeal control of Swedish word tones. Ann. Bull. RILP. 4:45-54.

<7> Howie, J.M. (1976) Acoustical studies of Mandarin vowels and tones. Cambridge: Cambridge University Press.

<8> Howie, J.M. (1974) On the domain of tone in Mandarin. Phonetica 30:129-148.

<9> Kratochvil, P. (1981) Tone and stress discrimination in normal Peking dialect speech. Proceedings of the second conference on Chinese Language Use. Canberra 1981.

<10> Kratochvil, P. (1984) Phonetic tone sandhi in Beijing dialect stage speech. Cahiers de Linguistique Asie Orientale XIII, 2:135-174.

<11> Kratochvil, P. (1985) Variable norms of tones in Beijing prosody. Cahiers de Linguistique Asie Orientale XIV, 2:153-174. Paris.

<12> Ohala, J.(1972) How is pitch lowered? Abstract in JASA 52,1 (part 1).

<13> Ohala, J. & H. Hirose (1970) The function of the sternohyoid muscle in speech. Ann. Bull. RILP 4:41-44.

<14> Sagart, L., P. Hallé, B. de Boysson-Bardies et C. Arabia-Guidet (1986) Tone production in Modern Standard Chinese: an electromyographic investigation. Cahiers de Linguistique Asie Orientale, XV, 2:205-221. Paris.