Ann. Bull. RILP. No. 14, 113-123 (1980)

LARYNGEAL CONTROL FOR INITIATION OF UTTERANCES
- A SIMULTANEOUS OBSERVATION OF GLOTTAL CONFIGURATION AND LARYNGEAL EMG -

Hajime Hirose, Masayuki Sawashima
and Hirohide Yoshioka

Introduction

In a recent issue of this Bulletin, we reported on detailed observation of laryngeal movements in terms of glottal abduction-adduction gesture with respect to the EMG patterns of the posterior cricoarytenoid (PCA) and interarytenoid (INT) for each of the utterance samples with various types of voiceless sounds and sound sequences in Japanese. Specifically, the data were illustrated and discussed for those utterance samples in which the sound or sound sequence in question was placed in the word medial position (Sawashima, Hirose and Yoshioka, 1978).

In the present report, observations will be confined to the conditions where similar utterance samples are in the absolute initial position, thus aiming to reexamine the notion of the transition from the "respiratory mode" to "speech mode" reported previously (Sawashima, Hirose, Ushijima and Niimi, 1975).

The mechanism of laryngeal control for different types of utterance initiation can be compared to that for different types of vocal attack. The term 'attack' usually refers to vocal initiation in singing, but may also be used to refer to utterance initiation of /CV/ sequences. Thus, Hirose and Gay (1973) examined three types of vocal attacks electromyographically and concluded that the three different types, i. e. breathy, soft and hard attacks, were characterized by coordinated actions of the abductor and adductor muscle groups of the larynx and, further, that these muscles worked in a reciprocal fashion for each type of vocal attack.

Most of these earlier results were based on a computer averaging technique for obtaining an averaged indication of overall muscle activity, and the comparison of EMG patterns with the glottal configuration in each utterance type was not necessarily straightforward. In the present study, simultaneous recordings of glottal image and pertinent muscle activities enabled us to attempt a more detailed comparison between the glottal movement and muscle activity patterns with respect to the initiation of utterances.

Procedures

Subjects

The subjects were two adult males (MS and HH), speakers of the Tokyo dialect.

Data recordings and processings

The glottal view of each subject was obtained through a fiberscope inserted through the nasal cavity and filmed by a 16mm movie camera at a rate of 50 frames per second, with simultaneous recordings of acoustic and EMG signals. The EMG recordings were made using hooked-wire electrodes inserted perorally to INT and PCA under indirect laryngoscopy. A more detailed description of instrumentations has been previously published (Sawashima, 1976).

The recorded EMG data were computer-processed for each token, in which the integrated EMG signals were smoothed by plotting running averages with a time window of 50 msec. Corresponding acoustic data were also processed in the same manner.

The film was projected on a semi-transparent screen to permit tracing of the film image for frame-by-frame analysis, and the distance between the tips of the vocal processes was measured as an indication of the glottal width.

Speech materials and experimental conditions

Eleven meaningful Japanese words were prepared as shown in Table I.

1. 異性 /iseH/
2. 畏敬 /ikeH/
3. 義兄 /gikeH/
4. 時勢 /ziseH/
5. 市営 /sieH/
6. 市税 /sizeH/
7. 気鋭 /kieH/
8. 奇形 /kikeH/
9. 吉慶 /kiQkeH/
10. 姿勢 /siseH/
11. 失政 /siQseH/

Table I

The high vowel /i/ between voiceless consonants in test words Nos. 8, 9, 10 and 11 was devoiced in both subjects. Consequently, these words have voiceless sound sequences consisting of voiceless consonants and the devoiced vowel [i̥]. No accent kernel is contained in the test words. Each word was placed in the absolute initial position of the test utterance and followed by a frame "——desu" (that is ——).

During the recordings, utterance samples were uttered successively with some intermediate pause, where the subjects were required to take a short breath at each time.

Result

In Fig. 1, integrated EMG curves of INT and PCA for a representative token of /iseH/ of each subject are shown together with the corresponding temporal change in the glottal width (GW) and the envelope of speech waves (AUDIO). An arrow in each graph indicates the onset of voicing for the

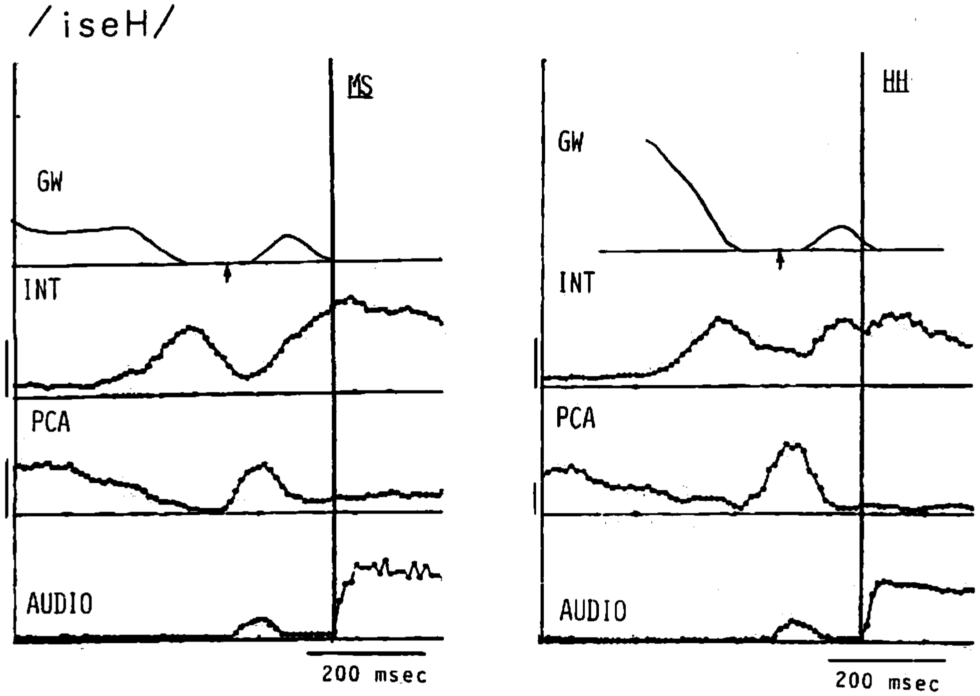

Fig. 1 Temporal course of the glottal width (GW), integrated and smoothed EMG curves of INT and PCA, and speech envelope (AUDIO) for /iseH/. The curves are aligned on the same time axis. The arrow indicates the onset of [i], and the vertical line indicates the voice onset of [i] after [s]. Calibration: 100μV.

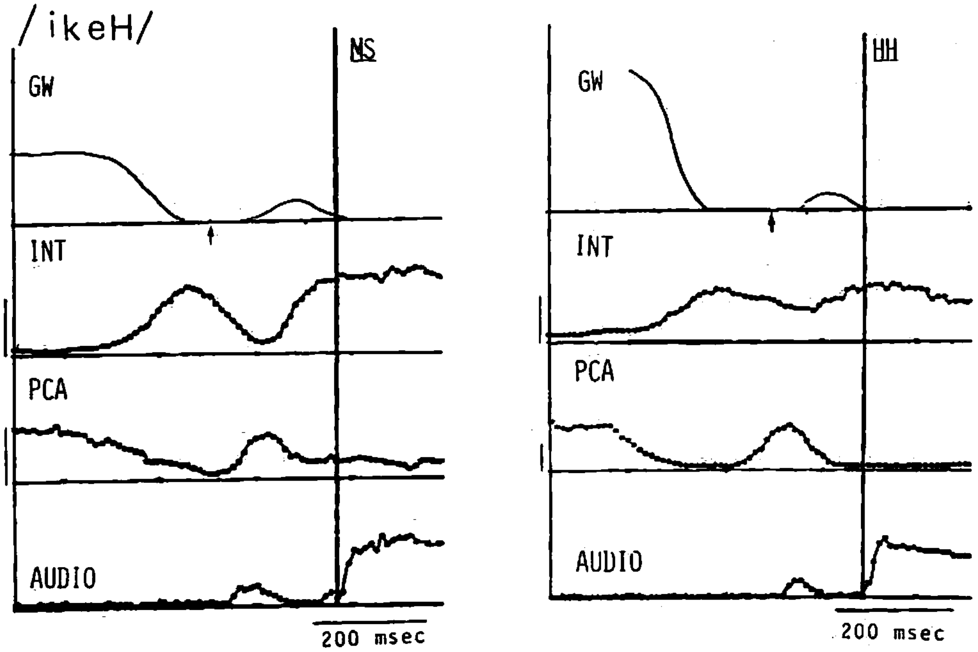

Fig. 2 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /ikeH/. The arrow indicates the onset of [i], and the vertical line indicates the voice onset of [e] after [k]. Calibration: 100μV.

initial [i], and a vertical line indicates the onset of the vowel [e] following the voiceless sound [s] in the test utterance. It can be seen in both subjects that the GW curve follows a smooth, gradual time course of glottal closure associated with a gradual increase in INT activity as well as with a gradual decrease in PCA, the latter being in nearly an inverse pattern of the INT curve. The INT curve reaches its peak slightly earlier than the PCA curve reaches the bottom of its trough. The timing of the INT peak roughly coincides with that of the approximation of the vocal processes preceding the initiation of utterance, while voicing of the initial vowel [i] starts with some delay. Actually, in these cases voicing starts even after PCA becomes activated for the production of the intervocalic /s/.

Very similar patterns are observed for /ikeH/ (Fig. 2), in which vertical lines indicate the onset of the vowel [e] after the stop [k].

Essentially similar patterns are also seen for utterance types /gikeH/ (Fig. 3) and /ziseH/ (Fig. 4), both of which are initiated by a voiced obstruent. It must be mentioned, however, that in HH, a complete approximation of the vocal processes is not obtained for initial voiced obstruents. For /gikeH/ in particular, the glottis tends to reopen after the initial gradual approximation where transient reactivation of PCA is noted. Presumably, the voiced /g/ was realized as a voiced palatal fricative in this case. In subject MS, the glottis closes approximately 80 msec before the initiation of voicing for these utterance samples. In this case, it seems that the glottis is already in the so-called phonatory or neutral position for voicing before the subglottal pressure builds up high enough for initiation of vocal fold vibration.

Utterance samples starting with a voiceless fricative /s/ followed by a vowel segment are shown in Fig. 5 (/sieH/) and Fig. 6 (/sizeH/). In both cases, the glottis remains fairly wide open until the very moment of the onset of [i] after [s], which is indicated by a vertical line in /sieH/, and, as a result, glottal closure is substantiated rather abruptly. Associated EMG patterns show a relatively steep increase in INT activity after the period of continuous suppression. The level of PCA activity stays relatively high when INT continues to be suppressed and then begins to decrease at the vicinity of the start of frication noise, which is indicated by an open triangle on the time axis. Also, a varying degree of "hump" is generally noted in the GW curve just before the abrupt narrowing of the glottis. The hump appears to become manifest when the glottis tends to reopen after the initial closing movement from the so-called respiratory phase or mode to speech mode, where a "dent" precedes the hump in the GW curve. In the case of /sieH/ of subject MS in particular, there is a marked hump peaking immediately after the start of frication noise. In this case, there is a temporary suppression of INT together with an instantaneous reactivation of PCA corresponding to the hump formation. In subject HH, on the other hand, the hump appears to become manifest simply by PCA activation without discernible transient suppression of INT.

A similar hump formation is also observed for utterance types /kieH/ (Fig. 7), /kikeH/ (Fig. 8) and /kiQkeH/ (Fig. 9), all of which start with a stop [k] followed by either voiced or unvoiced segments. The timing of the hump formation appears to be in the vicinity of the release of the initial stop [k], which is indicated by a solid triangle on the time axis. In /kikeH/ and /kiQkeH/, where the utterance starts with an unvoiced segment of [kik] or [kịkk], glottal closure after the hump is more gradual than in the case of /kieH/ or /sieH/. Particularly, in the case of /kiQkeH/ of subject MS, the glottis stays slightly open even for as long as 200 msec before the complete closure. In this case, both INT and PCA activities stay leveled correspondingly to the glottal condition. In subject HH, the time course of INT activation and PCA suppression for these utterance types are also extremely gradual.

The time courses of glottal closure for utterance types /siseH/ (Fig. 10) and /siQseH/ (Fig. 11) starting with an unvoiced segment [sịs] or [sịss] are similar to those for /kikeH/ or /kiQkeH/. Namely, a relatively long open glottis period continues after the onset of frication noise (indicated by an open triangle) before the voice onset of vowel [e], which is marked with a vertical line. The temporal patterns of the INT and PCA curves are again in an inverse relationship to each other, and very gradual activation of INT associated with gradual suppression of PCA is the basic pattern of muscle control for the glottal dynamics in these samples as well. These temporal courses of the GW curve and EMG patterns obtained in the present experiment are quite comparable to those observed in the closing phase of the glottis in the production of the same utterance samples /siseH/ and /siQseH/ placed in the word medial position. In other words, the prephonatory glottal dynamics in these cases appear to be realized similarly in the "speech mode."

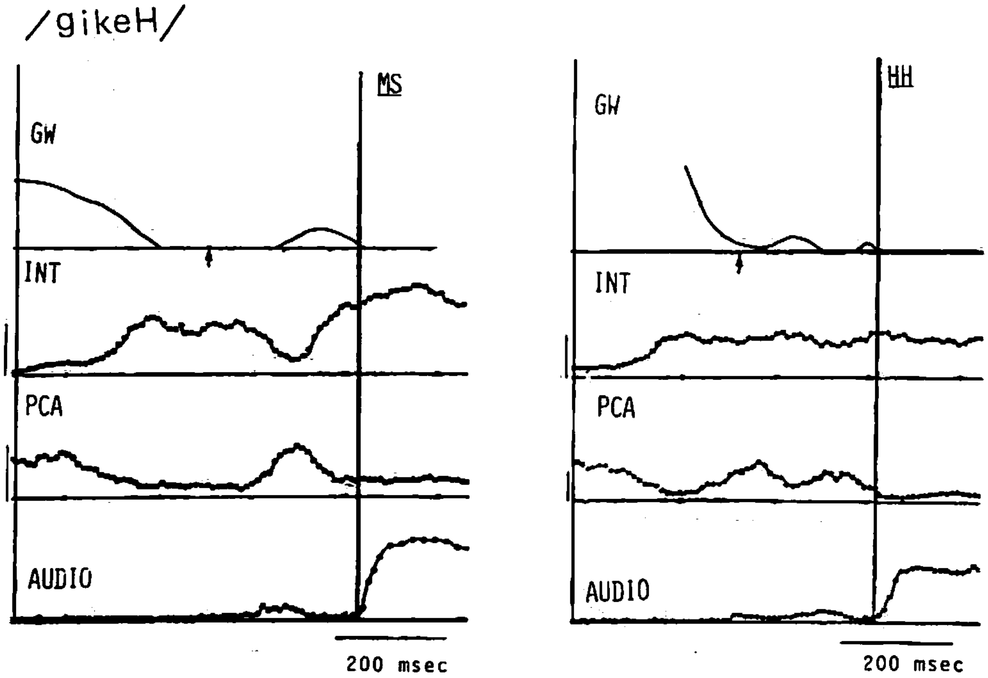

Fig. 3 Temporal xourse of GW, EMG curves of INT and PCA, and AUDIO for /gikeH/. The arrow indicates the onset of [g], and the vertical line indicates the voice onset of [e] after [k]. Calibration: 100μV.

/ziseH/

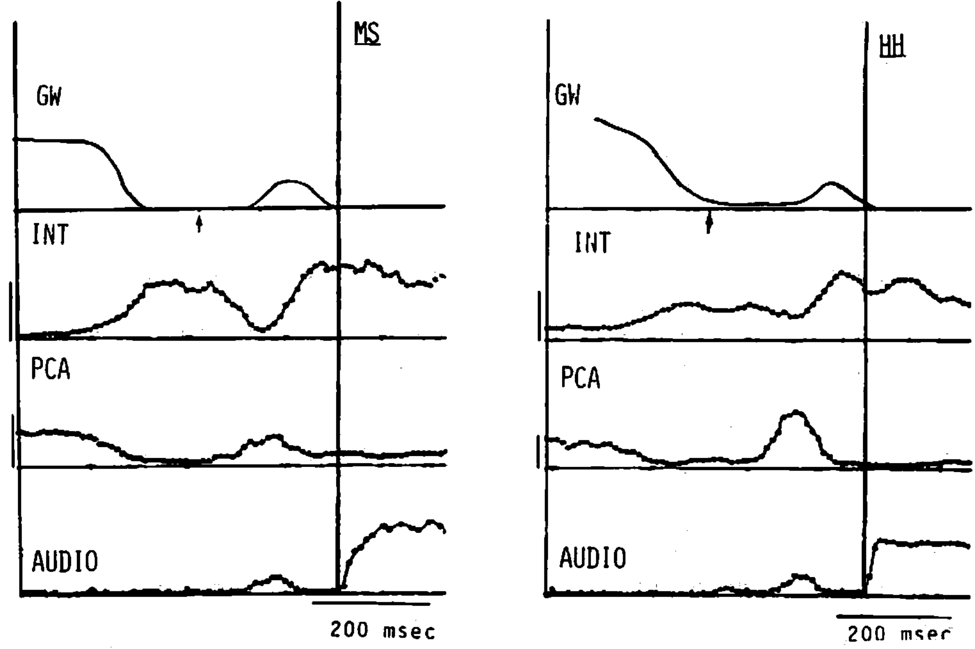

Fig. 4 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /ziseH/. The arrow indicates the onset of [x], and the vertical line indicates the voice onset of [e] after [s]. Calibration: 100 μV.

/si'eH/

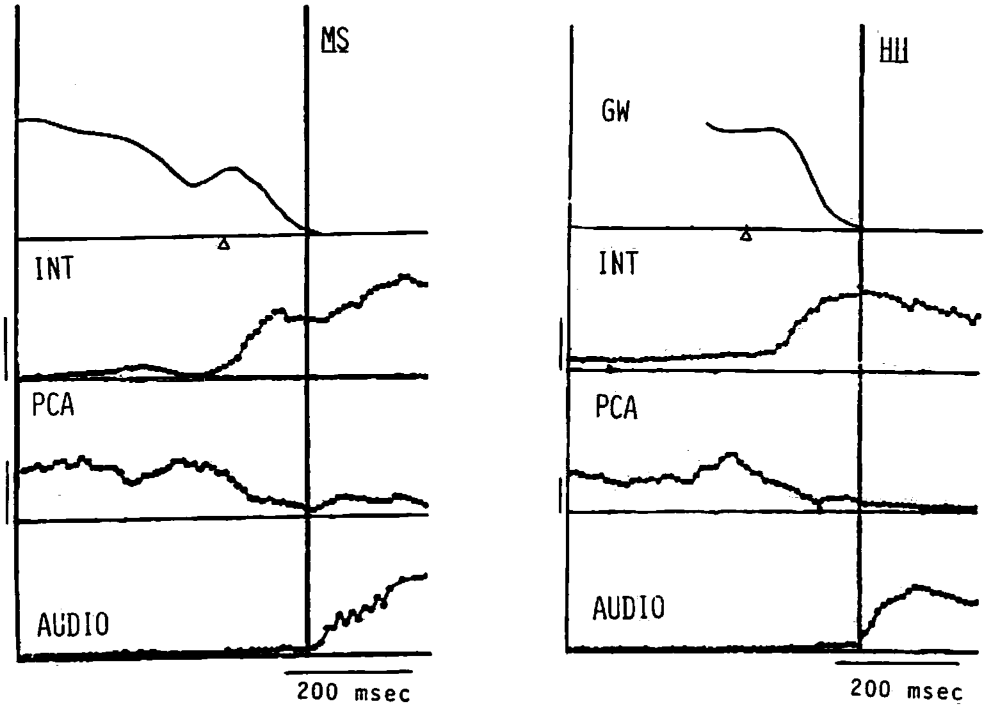

Fig. 5 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /si'eH/. The open triangle indicates the onset of frication noise, and the vertical line indicates the voice onset of [i] after [s]. Calibration: 100 μV.

/sizeH/

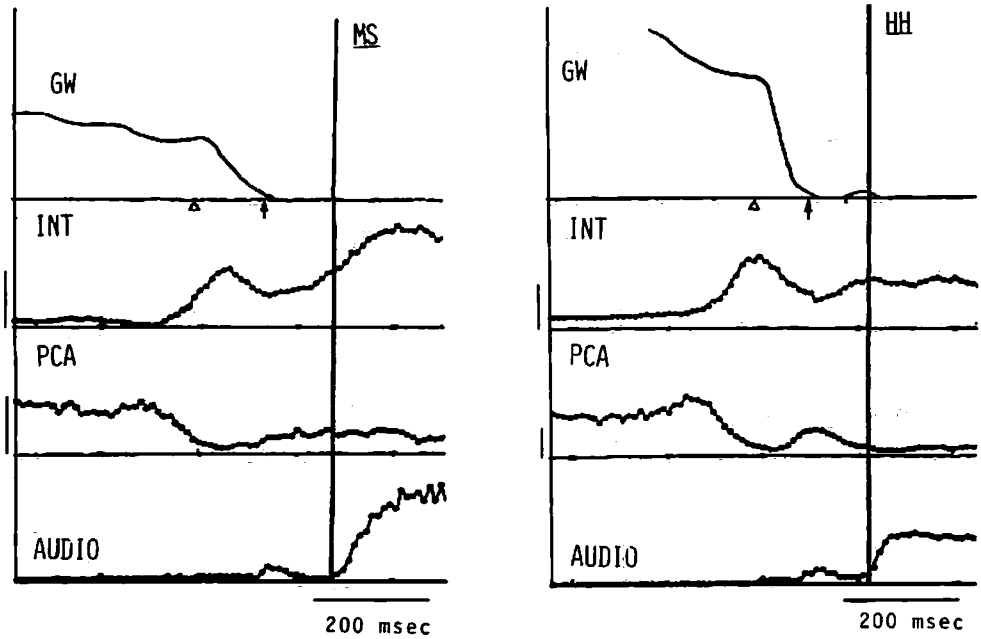

Fig. 6 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /sizeH/. The arrow indicates the onset of [i], the open triangle indicates the onset of frication noise, and the vertical line indicates the voice onset of [e] after [z]. Calibration: 100μV.

/ki'eH/

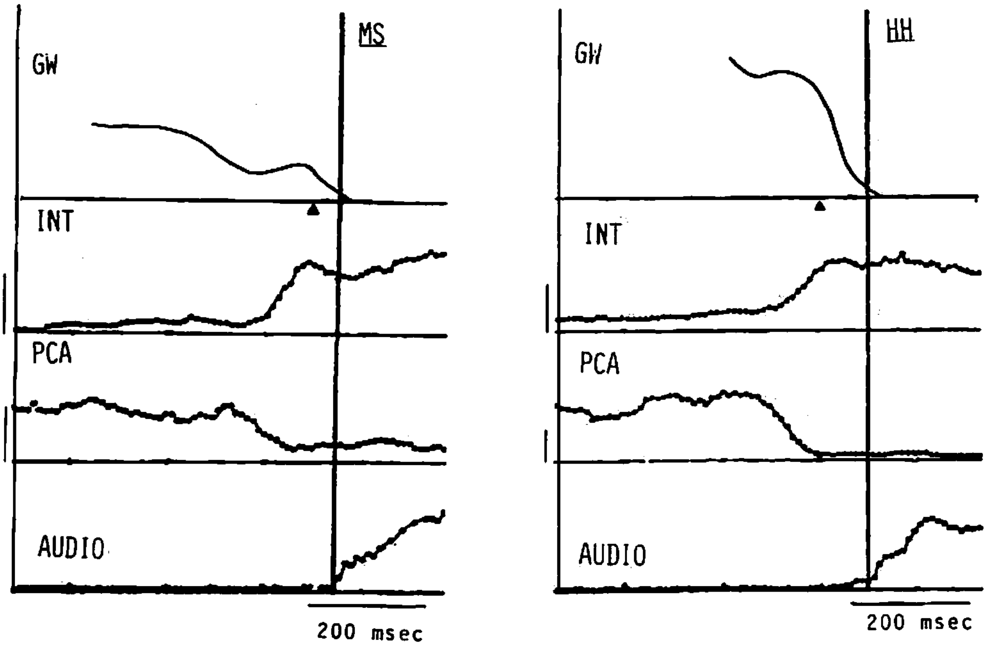

Fig. 7 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /ki'eH/. The filled triangle indicates the release of closure before the vowel [i], and the vertical line indicates the voice onset of [i]. Calibration: 100 μV.

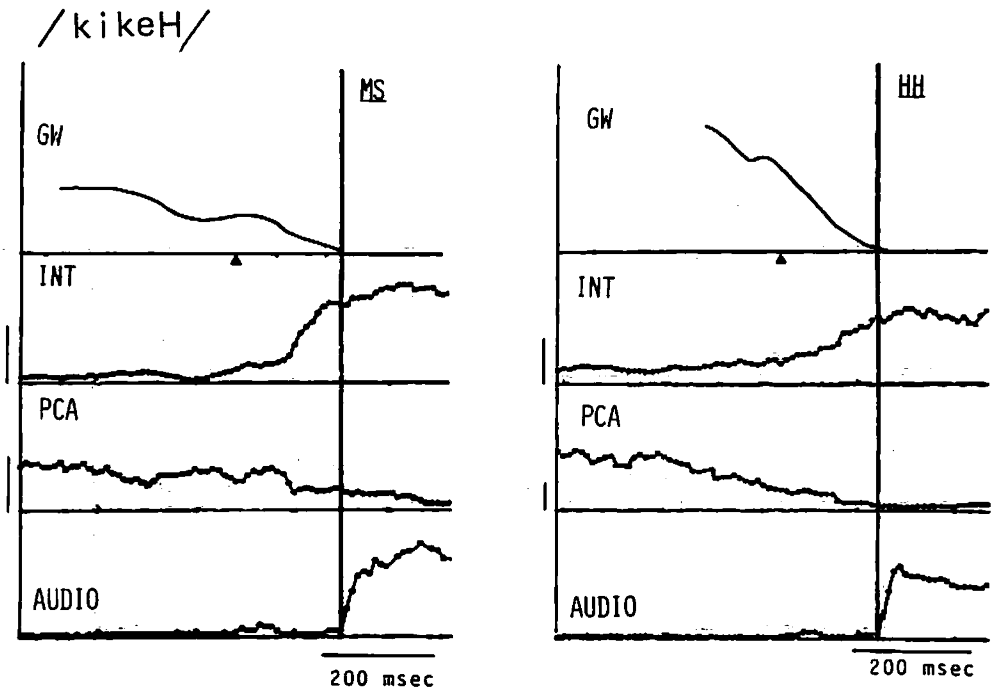

Fig. 8 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /kikeH/. The filled triangle indicates the release of the initial stop [k], and the vertical line indicates the voice onset of [e] after [ki̥k]. Calibration: 100 μV.

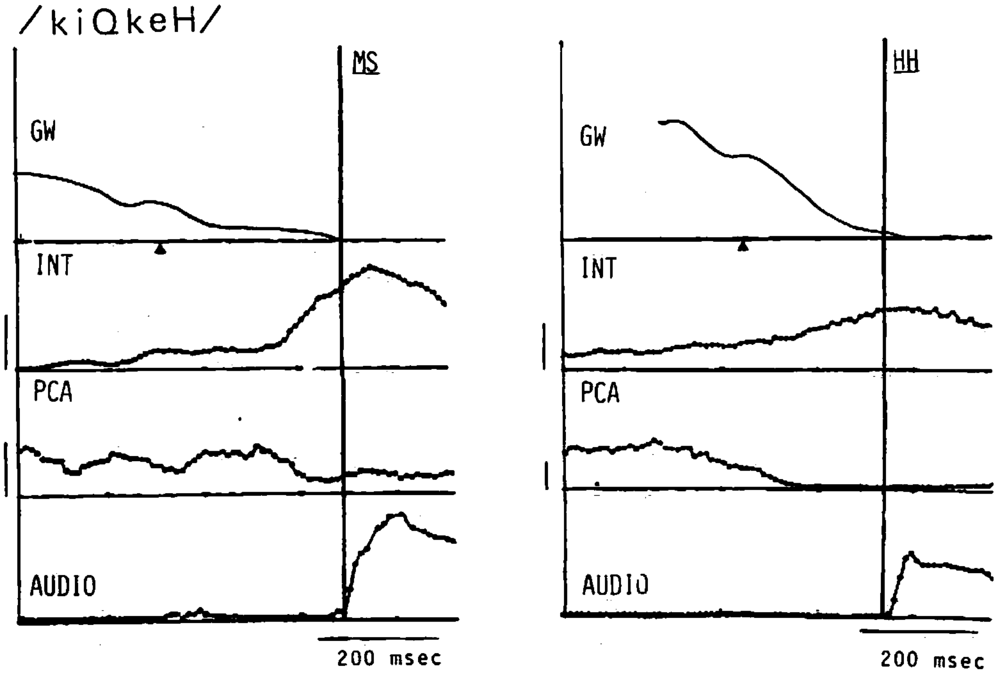

Fig. 9 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /kiQkeH/. The filled triangle indicates the release of the initial stop [k], and the vertical line indicates the voice onset of [e] after [ki̥k]. Calibration: 100 μV.

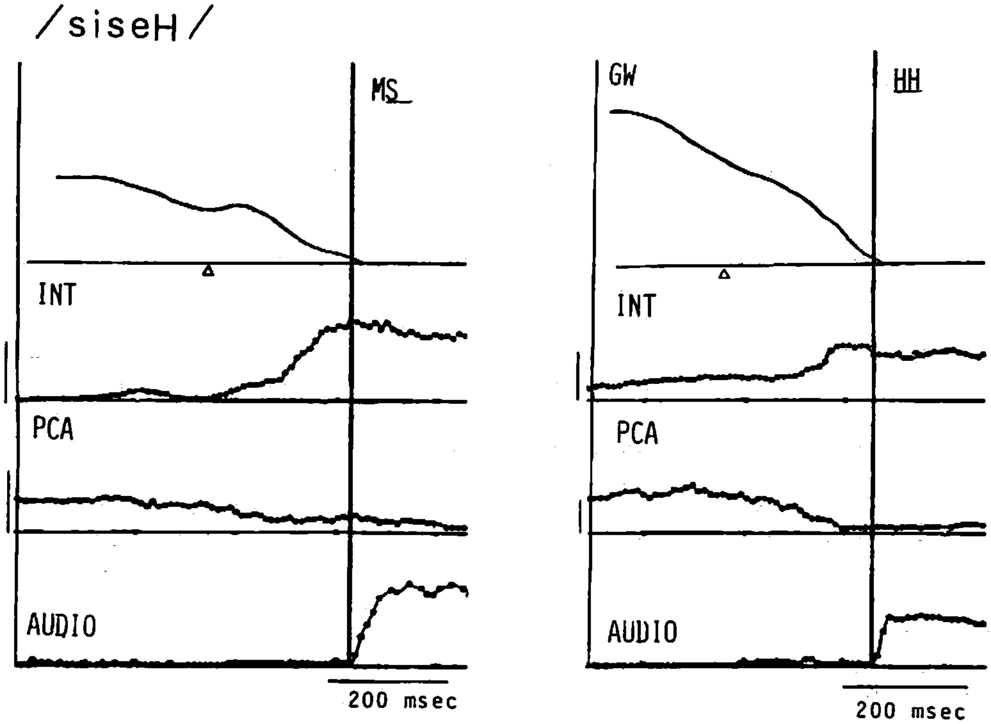

Fig. 10 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /siseH/. The open triangle indicates the onset of frication noise, and the vertical line indicates the voice onset of [e] after [si̥s]. Calibration: 100 μV.

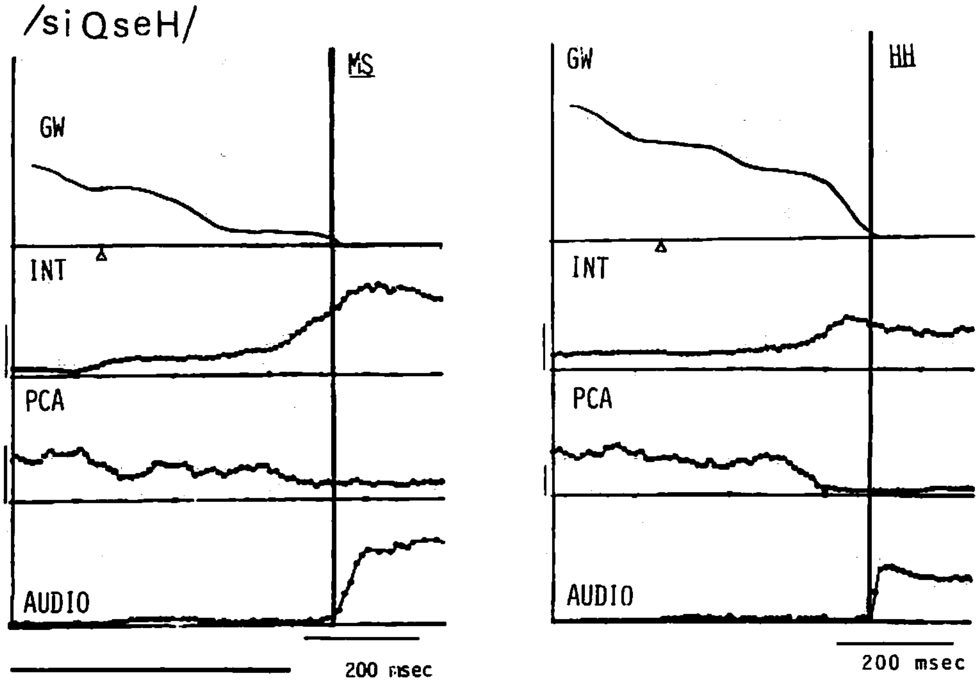

Fig. 11 Temporal course of GW, EMG curves of INT and PCA, and AUDIO for /siQseH/. The open triangle indicates the onset of frication noise, and the vertical line indicates the voice onset of [e] after [si̥s]. Calibration: 100 μV.

Comment

It is confirmed in the present study that there is a voiced-voiceless distinction in the laryngeal adjustments for initiation of utterances as reported previously. The distinction is characterized by a temporal delay of the glottal closure for the presence of voiceless sound or sound sequences in the absolute initial position. In some cases, a hump formation is noted in the time course of glottal closure which can be taken as an evidence of a volitional abduction gesture for maintaining an adequate glottal width for unvoiced condition. Thus, the glottal gesture for initiation with voiceless sound(s) resembles that for breathy attack. If the utterance starts with voiced sound(s), the glottis closes gradually from the widely open condition for prephonatory inspiration, regardless of the presence or absence of a glottal stop gesture at the initiation of voicing.

It is also confirmed that these glottal gestures for initiation of utterances are realized by coordinated activities of the abductor (PCA) and adductor (INT) of the vocal fold, both of which are activated generally in a reciprocal fashion. However, as mentioned in the previous report in which the same two subjects MS and HH were examined in a different experimental session, there is a certain individual difference in the pattern of muscle control for a given specification of laryngeal articulation. In subject HH, a fine adjustment of the glottal condition in terms of vocal fold abduction is mainly achieved by PCA activity, while INT simply shows overall suppression for unvoicedness. In subject MS, on the other hand, INT appears to participate more actively in a fine glottal adjustment for unvoiced condition in that transient or continuous INT suppression is always associated with PCA activation. These findings are quite comparable to those found previously with respect to the glottal adjustment for consonant production in the word medial position.

In our previous data dealing with glottal adjustment in the absolute initial position using the same subjects, PCA activity in subject MS appeared to be rather indifferent to the control of the glottal width, and it was assumed that this particular subject used the so-called "respiratory mode" for initiation of utterances. As the present data indicate, however, PCA activity in reverse pattern to that of INT is also quite relevant to the glottal adjustment for utterance initiation in subject MS as well. Presumably, the dynamic range of recorded PCA activity in the previous experiment was not high enough to reveal a clear reciprocity of PCA with INT in the word initial position. It may also be plausible that the averaging procedure resulted in an unavoidable time smearing particularly in the vicinity of the initiation of utterances, where even intrasubject variations might take place for each utterance. In any case, it seems tenable to conclude that both subjects use PCA and INT in the transition from the respiratory phase to the initiation of utterances with a slightly different individual strategy of laryngeal muscle control.

Acknowledgement

This study was supported in part by the Grant in Aid for Scientific Research (No. 448322, No. 410207, No. 337040), Ministry of Education, Science and Culture.

References

Hirose, H. and Gay, T. (1972): The activity of the intrinsic laryngeal muscles in voicing control: an electromyographic study. Phonetica 25, 140-164.

Hirose H. and Gay, T. (1973): Laryngeal control in vocal attack. Folia phoniat. 25, 203-213.

Hirose, H. (1976): Posterior cricoarytenoid as a speech muscle. Ann. Otol. Rhinol. Laryng. 85, 334-343.

Sawashima, M. (1976): Current instrumentation and techniques for observing speech organs. Technocrat 9, No. 4, 19-26.

Sawashima, M., Hirose, H., Ushijima, T. and Niimi, S. (1975): Laryngeal control in Japanese consonants, with special reference to those in the utterance initial position. Ann. Bull. RILP., 9, 21-26.

Sawashima, M., Hirose, H. and Yoshioka, H. (1978): Abductor (PCA) and adductor (INT) muscles of the larynx in voiceless sound production. Ann. Bull. RILP 12, 53-60.