Ann. Bull. RILP, No. 9, 21-26 (1975)

LARYNGEAL CONTROL IN JAPANESE CONSONANTS, WITH SPECIAL REFERENCE TO THOSE IN UTTERANCE-INITIAL POSITION

M. Sawashima, H. Hirose, T. Ushijima and S. Niimi

Fiberscopic observation of the larynx during articulations of Japanese consonants has revealed that there is an opening of the glottis for voiceless sounds. The degree of the opening varies with different phonemes and phonological environments. [1] In contrast, there is no appreciable glottal opening gesture for voiced sounds.

Electromyography of the intrinsic laryngeal muscles during speech has also revealed that the voiced vs. voiceless distinction at the level of the larynx is accomplished by a clear reciprocal activity pattern in the abductor muscle (posterior cricoarytenoid) and the adductor muscle (interarytenoid). [2] An increase in the activity of the posterior cricoarytenoid (PCA) muscle, associated with a decrease in the activity of the interarytenoid (INT) muscle, is observed for voiceless sounds; while for the voiced sounds a decrease in the activity of PCA is associated with an increase in the activity of INT.

Observation of the EMG pattern of PCA in correspondence with the fiberscopic data in some selected samples has further revealed that there is a positive correlation between the peak value of PCA activity and the value of the maximum glottal width during articulations of the voiceless consonants, and that the time course of the glottal opening and closing is best represented by the temporal pattern of PCA activity, among the intrinsic muscles. [3), 4)]

All of the data just mentioned were obtained for those sounds which are positioned utterance-medially, even though some of them are in word-initial position. Thus, these data should be assumed to describe adjustments of the larynx when it is in the so-called "speech mode".

In the present experiment, we were interested in whether there is a voiced vs. voiceless distinction in the laryngeal adjustments utterance-initially, i. e. at the transition from the "respiratory mode" to the "speech mode"; and if so, how it appears.

Experimental Procedures

Two adult males (M. S. and H. H.), speakers of the Tokyo dialect, served as subjects. A set of meaningful words was prepared for each of the subjects. The words are:

Subj. MS

制定	[se : te:]	定性	[te : se :]	軽々	[ke : ke :]
ペーペー	[pe : pe:]	米兵	[be : he :]	税制	[ze : se :]
設定	[sette :]	設計	[sekke :]	閲兵	[eppe :]

Subj. HH

制定 [se : te:] 定性 [te : se :] 命名 [me : me :]
ペーペー [pe : pe:] 米兵 [be : he :] 英米 [e : be :]
税制 [ze : se :] 泥々 [de : de :] 営々 [e : e :]
平生 [he : ze :]

In both sets of test words, voiceless consonants and voiced sounds including the vowel [e] are positioned word-initially, and voiced and voiceless consonants word-medially. No accent kernel is attached to these words in the Tokyo dialect. The subject repeated the list ten times in random order, each word being followed by a frame "--- to yuu". (We say - - -.) The speaking rate and the vocal intensity as well as the vocal pitch were kept as constant as possible for each subject within his habitual range. The subjects were also instructed to place a long enough interval between each utterance to return to the respiratory position and to take one breath (inspiration) before utterance initiation.

A fiberscope was inserted through the nose of the subject and the glottal view was photographed by a 16 mm cine camera at a frame rate of 50 per second with simultaneous recording of speech signals. Frame-by-frame signals of the cine camera and timing marks were also recorded on one channel of a stereo-audio tape recorder, speech signals being recorded on the other channel. On the oscillographic trace of these signals, each of the film frames can be matched with the time course of the speech signals with an accuracy of approximately 10 msec. This system is described elsewhere in this issue of the Annual Bulletin. 5)

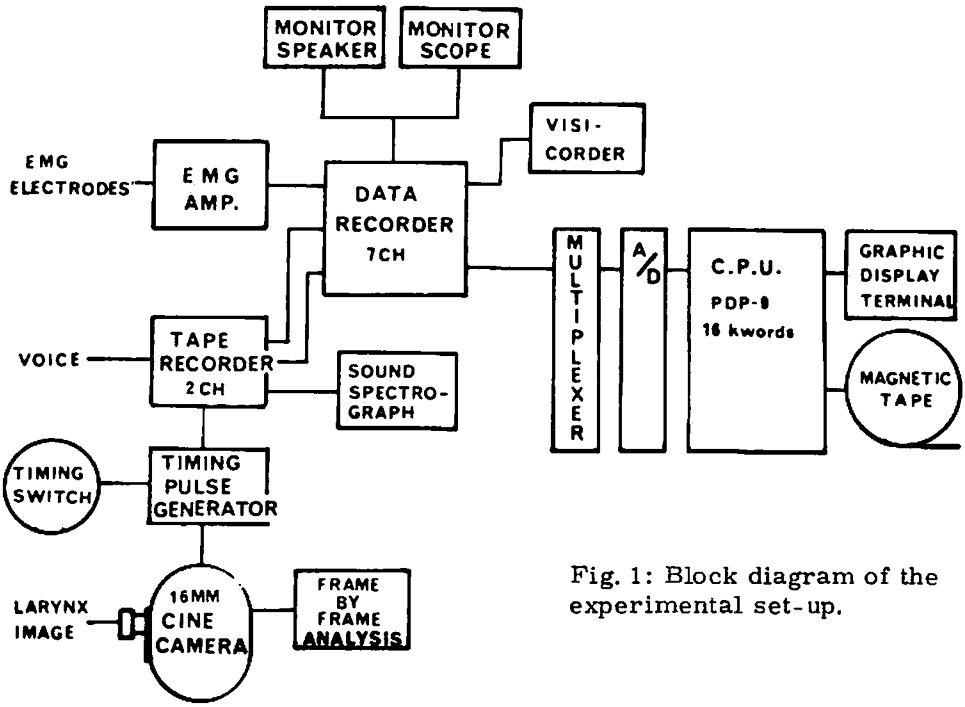

Fig. 1: Block diagram of the experimental set-up.

On the film image, the distance between the vocal processes of the arytenoid cartilages was measured frame-by-frame for the time period to be examined.

Simultaneously with filming and sound recording, EMG of the intrinsic laryngeal muscles were also recorded. Hooked wire electrodes were introduced into the vocalis (VOC) muscle through the neck, and into the PCA and INT muscles through the mouth. EMG signals were recorded on a multi-channel FM tape recorder along with the speech and camera signals recorded on other channels. EMG signals were later processed by a computer and an averaged EMG pattern for each muscle was obtained.

A block diagram of the experimental set-up is given in Fig. 1, which is identical to the one appearing in the last issue of the Annual Bulletin. [4]

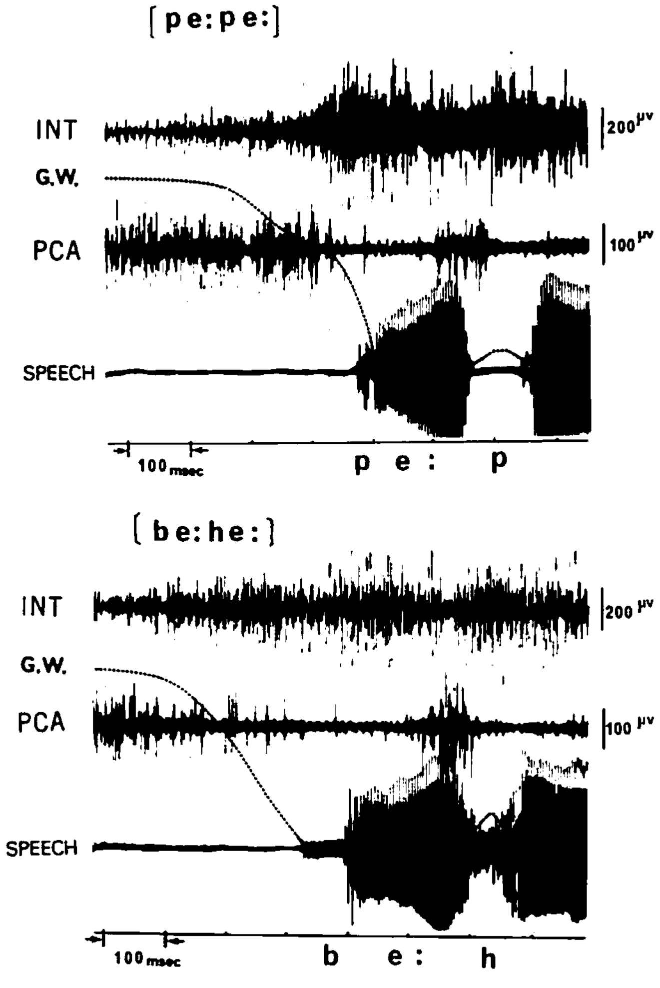

Fig. 2: Raw EMG data of PCA and INT compared with the time course of the glottal width (G.W.) for a /p/ vs. /b/ pair at the utterance initial position.

Results and Remarks

In subject HH, EMG of VOC was omitted from the data because of contamination by artefacts. Figure 2 displays examples of single utterances for [pe: pe:] and [be: he] by subject HH. For each utterance, the temporal change in the glottal width (G. W.) is superimposed on the same time axis of the traces of EMG and speech signals. It is noted that there is a clear difference in the time course of glottal adduction as well as muscle activities at utterance initiation between [pe: pe:] and [be: he:]. Fig. 3 illustrates the averaged time pattern of the glottal width and EMG of the two subjects. In each subject, three different types of utterances are contrasted, namely voiceless consonant [p], voiced consonant [b], and vowel [e] in utterance-initial position respectively. For averaging, EMG signals and glottal width curves are lined up at the vowel onset of the first syllable which is indicated as a vertical line in the graph.

In subject HH, when there is a vowel in utterance-initial position, the glottis begins to close gradually with a gradual decrease in the activity of PCA. The closing movement of the glottis becomes accelerated when the activity of the INT begins to increase. In general, the time course of the vocal fold adduction for utterance initiation is very smooth with gradual replacement of PCA activity by INT activity. This pattern is also observed in the glottal adjustment when [b] or one of the other voiced consonants is in utterance-initiation position. In the case of the voiceless consonant positioned utterance-initially, for example [p] as shown in the figure, the glottis stays wide open until less than 100 msec before the vowel onset. The activity of PCA is also maintained at a very high level. The glottis then closes very rapidly with an abrupt decrease in PCA activity and increase in INT activity. The onset of INT activity is markedly delayed as compared with that for the voiced sound.

In subject MS, the general pattern of the glottal adjustments is essentially the same as in the case of HH, when the voiced sound is placed utterance-initially, although the timing of the cessation of PCA activity appears to be relatively sooner. The activity of the vocalis (VOC) muscle begins to increase at almost the same time as that of INT. In the case of the voiceless consonant [p], there is also a delay in the closing action of the glottis but the pattern of the adjustments is quite different from that in HH. A gradual decrease in the glottal width begins as early as that for voiced sounds, and PCA activity also decreases with a time pattern nearly identical to that of the voiced sounds. There is a marked delay, as was seen in HH, in the onset of INT and VOC activity. Consequently there is a time period of nearly 100 msec where neither the abductor muscle nor the adductor is active. Corresponding to this time period there is a considerable slow-down of the closing movement of the glottis, before it is accelerated by the activity of the adductor muscle. This particular pattern of the glottal adjustments is also observable for other voiceless consonants (in the utterance-initial position) in this subject.

From the results it is clear that for both subjects there is a voiced vs. voiceless distinction in the laryngeal adjustments for initiation of utterances. The distinction is characterized by a temporal delay of the vocal fold adduction for the voiceless sound, which corresponds to the vocal fold abduction for the voiceless sound in the utterance-medial environment.

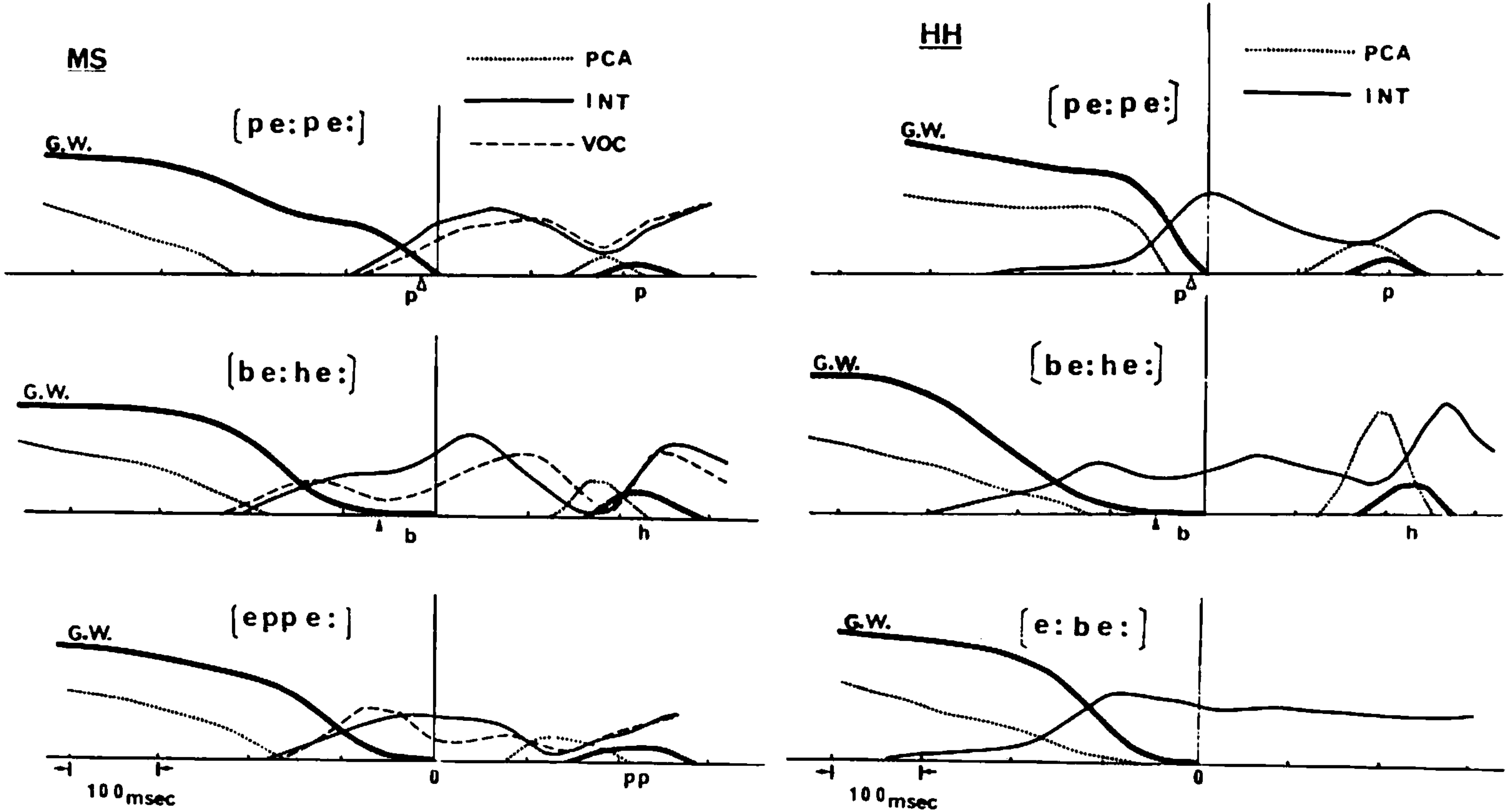

Fig. 3: Comparison between the time course of the glottal width (G.W.) and the averaged EMG curves for subjects HH and MS. Δ indicates the release of [p] and ▲ the onset of voicing of [b]. The onset of vowel in the first mora is taken as the line-up point for averaging.

One should note here, however, that the temporal delay of the vocal fold adduction in the two subjects is realized differently in both subjects. In subject HH, there is persistent high activity of PCA associated with delayed onset of INT activity. In subject MS, on the other hand, there is a delay in the onset of INT and VOC activities without any particular activation of PCA. In other words, the laryngeal specification for voicelessness of the utterance-initial sound is actualized by an active mode involving contraction of the abductor muscle in subject HH, but in subject MS, by a passive mode with relaxation of both the abductor and adductor muscles. This kind of significant subject-to-subject variation in the laryngeal muscle control for a given specification probably may occur only in limited situations such as the transition from the "respiratory mode" to the "speech mode" (of the larynx), at the initiation of an utterance. In Fig. 3, glottal adjustments for the word-medial voiceless consonants are also illustrated. For both subjects, there is an increase in the activity of PCA associated with a suppression of the activities of INT and VOC. Among these muscles, the temporal pattern of PCA activity appears to best represent the temporal change in the glottal width. This finding is in good agreement with our previous reports. There seems to be more suppression of the adductor activity in subject MS than in HH, but the glottal opening appears to be primarily controlled by PCA activity as in the case of HH. Thus, the laryngeal control for a given specification appears to become quite uniform, with less individual variation, when the larynx is set in the "speech mode".

It is well known in the physiology of the larynx that the vocal fold at rest stays in an intermediate position which is nearly midway between fully abducted and adducted positions. Activities of both the adductor and abductor muscles are minimum and the glottis is considerably open. This position is usually observed during quiet respiration and is regarded as a characteristic status of the larynx in the "respiratory mode".

One can assume that in the transition from the "respiratory mode" to the "speech mode", the glottal opening for the utterance-initial voiceless sound may be actualized in either the "respiratory mode" or the "speech mode", and the former is taken by subject MS and the latter by subject HH.

References

1. M. Sawashima and S. Niimi (1974): "Laryngeal conditions in Articulations of Japanese Voiceless Consonants." Ann. Bull. RILP, No. 8 13-18.

2. H. Hirose et al. (1972): "The Activity of the Intrinsic Laryngeal Muscles in Voicing Control - An Electromyographic Study. Phonetica 25:140-164.

3. H. Hirose and T. Ushijima (1974): "The Function of the Posterior Cricoarytenoid in Speech Articulation." Status Report on Speech Research (Haskins Laboratories) SR 37/38: 99-107.

4. S. Niimi and M. Sawashima (1974): "A Preliminary Study on the Simultaneous Recording of Laryngeal Muscle Activities and the Glottal Shape During Speech Utterance." Ann. Bull. RILP. No. 8:19-22.

5. S. Sekimoto et al. (1975): "Synchronization of a Film Image with a Speech Signal." Ann. Bull RILP No. 9: