AN ELECTROMYOGRAPHIC STUDY OF THE ACTIVITY OF THE LARYNGEAL MUSCLES DURING SPEECH UTTERANCES*

Hajime Hirose, Zyun'ici Simada, and Osamu Fujimura

The aim of the present study is to examine the electromyographic activities of selected laryngeal muscles which are considered to serve as adductors of the vocal cords during speech utterances. In the experiments to be discussed here, the thyroarytenoid muscle and the lateral cricoarytenoid (lateralis) muscle were examined. Strictly speaking, the thyroarytenoid muscle consists of multiple anatomical portions. In the present study, however, the most medially situated portion (the internal thyroarytenoid or the vocalis muscle) was examined.

The vocalis muscle arises from the inner surface of the thyroid cartilage at the angle formed by the junction of the laminae, extends backward and attaches to the arytenoid cartilage mainly at the vocal process. It pulls the arytenoid cartilage forward at the point of attachment and consequently rotates the cartilage medially. For this reason, it is generally believed to have an adducting effect in addition to shortening and tensing effects on the vocal cord.

The lateralis muscle arises from the upper border and the outer surface of the cricoid arch, extends upward and backward, attaches to the anterior surface of the muscular process of the arytenoid cartilage. By its contraction, it pulls the muscular process forward and downward resulting in an adduction of the vocal cord. Figure 1 presents a schematic drawing of the anatomical locations of the vocalis and the lateralis muscles.

Recent technical developments in electromyography have made it possible to obtain data on the laryngeal muscle activity during speech and

---

* A modified English version of a paper read at the Semiannual Meeting of the Acoustical Society of Japan, May 7, 1970.

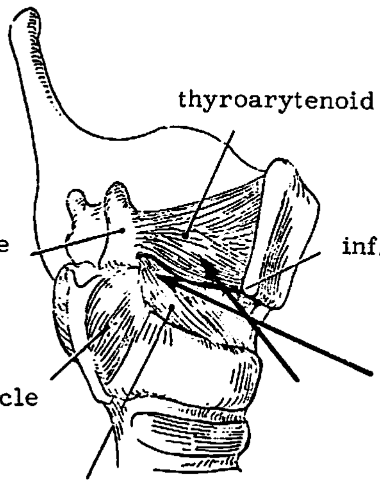


Figure 1. A lateral view of the larynx with the right ala of the thyroid cartilage removed. Arrows indicate the direction of insertion of the needles to the vocalis muscle and to the lateral cricoarytenoid muscle, respectively. (This figure was taken from Hirano and Ohala* with slight modifications.)

singing particulary effectively by the use of the double-ended hooked-wire electrodes inserted to the pertinent muscle through a percutaneous approach.* In our experiments, a pair of insulated copper wires are employed as the electrodes, the outer diameters of which are approximately 80 micron. The pieces of wire are threaded through a 27-gauge hypodermic needle (0.4 mm in outer diameter and 2.0 cm in length) and the needle carrying the electrodes is inserted percutaneously either to the vocalis muscle through the crico-thyroid membrane or to the lateralis muscle through the cricothyroid muscle. The arrows in Figure 1 represent the directions of the needle insertions in obtaining access to the indicated muscles. When the insertion is made, the location of the pair of electrodes is tested by observing the electromyographic signals induced by appropriate gestures that have been proved pertinent for the contraction of the selected muscle. If the validity of location is ascertained, the needle is withdrawn, leaving the electrodes hooked

* M. Hirano and J. Ohala, "Use of Hooked-Wire Electrodes for Electromyography of the Intric Laryngeal Muscles," J. of Speech and Hearing Research, Vol. 12, 362-373 (1969).

into the desired portion in the muscle. Once the electrodes have been thus fixed, the subject can talk in a natural manner with practically no discomfort. After the completion of the experimental procedure, the wires can be removed with a light tug.

The electromyographic and speech signals are recorded by an FM magnetic tape recorder. We employ a PDP-9 computer for processing the electromyographic data. The magnetic tape is played back and the recorded signals are fed to the computer through an A-D converter. The electromyographic signal is sampled every 250 μs and digitized into 6-bit levels. The absolute values are taken and then integrated and smoothed over a range of 10 ms. For a selected 10 out of the 15 recorded utterances for each item, the integrated values as functions of time are added together with reference to the point in the time axis representing a predetermined speech event observed in the speech signal of the test word.

In the present study, in order to simultaneously observe the conditions of the golttis while recording the electromyographic signals, a fiberscope* was inserted through the nose of the subject and a 16-mm motion picture was taken at 24 frames per second. Synchronizing time marks were provided both on the film and on the magnetic sound tape in the form of image-blanking DC signals with a sharp step-onset wave form.

Selected Japanese words as shown in Table 1 were used as the test words. The subject uttered each test word repeatedly 15 times in isolation, with a short pause between utterances.

It has been noted that high vowels between voiceless consonants are often devoiced in many Japanese dialects. In the test words, the vowel /i/ in /sitee/, /sisee/, /sihee/, /siQtee/ and /siQsee/ can be either voiced or devoiced. In the present experiment, one subject, with experience in phonetic studies first uttered these test words with and without devoicing in two separate series, series A and B, respectively. The subject then uttered each of the test words as natural utterances of the word in isolation, without paying

* M. Sawashima, "Movement of the Larynx in Articulation of Japanese Consonants," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo), No. 2, 11-20 (1968).

sitee	(私邸)	zitee	(自邸)
sisee	(姿勢)	zisee	(時勢)
sihee	(私兵)	zihee	(時弊)
si'ee	(私営)	zi'ee	(自営)
sidee	(紫泥)		
siQtee	(失邸)	ziQtee	(実弟)
siQsee	(失政)	ziQsee	(実勢)

Table 1. List of the test words used in the present experiment.

any attention to making the samples consistent in respect to the devoicing feature (series C). The recorded samples were later separated into two subsets for processing by the visual inspection of the acoustic signal on the visicorder trace.

Figure 2 exemplifies two samples of the raw data of the electromyographic and acoustic signals representing two different utterances, a repeated phonation of the vowel [e] in high-pitched chest voice of about 200 Hz, and a phonation of a neutral vowel with glottal stops. It is clearly seen in the figure that the activity of the vocalis muscle (see the lower trace) starts considerably earlier than the onset of the acoustic effect of voicing. In these examples with glottal stops, the start of the heavily activated electrical discharges leads the onset of the sound signal by approximately 400 ms,

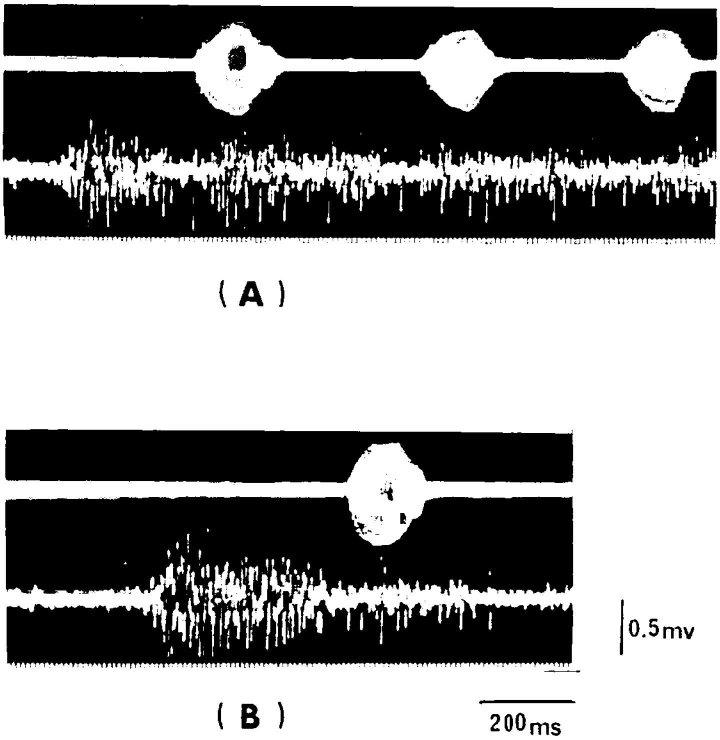


Figure 2. Examples of raw electromyographic and acoustic data of the vocalis muscle for repeated phonation of the vowel [e] in high-pitched chest voice (A), and for a neutral vowel containing glottal stops (B). Upper traces:acoustic signal; lower traces: electromyogram.

presumably indicating a tight glottal closure in preparation for the glottal plosion. It should also be mentioned here that the activity of the vocalis muscle is seen to be suppressed somewhat prior to (about 50 ms before) the voice onset until a sharp impulse corresponding exactly to the initiation of the vocalized vibration marks the start of muscular activities that are seen throughout the repeated occurrences of voicing.

Figure 3 shows a pair of sound spectrograms of the test word /sisee/ uttered with and without devoicing in the separate series A and B, respectively. Short vertical bars under the spectrograms demarcate time intervals for the successive frames of the motion picture. The arrow at the upper corner of each spectrogram indicates the start of the blanking signal which was placed immediately after the completion of the utterance for the synchronization of the sound record and the film frames. The numbers attached indicate the frame numbers counted back from the synchronization mark and correspond to those in Figure 4 and 5 (see infra). It may be mentioned that in the upper spectrogram, even though the vowel [i] is completely voiceless in the devoiced sample, we can identify rather clearly the manifestation

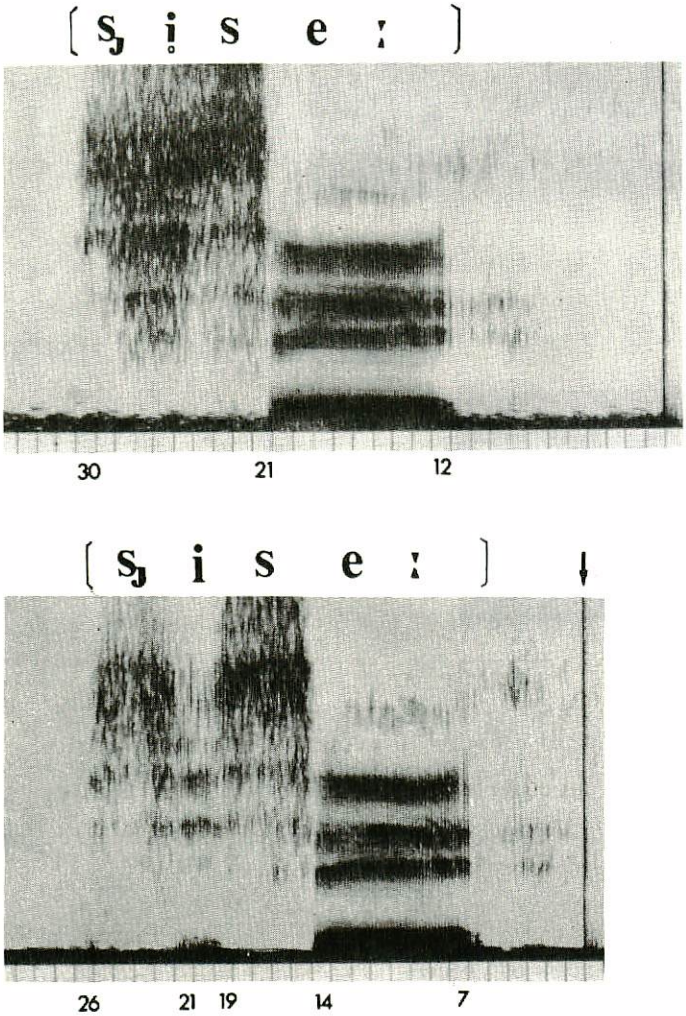


Figure 3. A pair of sound spectrograms of the test word /sisee/ uttered with and without devoicing in the separate series. The numbers attached correspond to those in Figures 4 and 5.

of the vowel as a time segment with distinctly intensified F_3, F_4 and F_5 regions (but with almost the same formant frequencies) in contrast to the adjacent fricative segments, which in turn show more concentrated components representing a higher resonance. This fact suggests that the location of the turbulence source differed between the consonantal segment and the voiceless vowel segment. We may thus assume that the relative airflow impedances at the vocal tract constrictions, i. e., the glottal and palatal constrictions, have been appreciably different for the two segments, even though the tract configuration had some clear similarity, perhaps with different absolute areas but with quite similar proportions, length, and shape near the constriction. As a matter of fact, it would not be necessary, for the interpretation of the above mentioned acoustic observation, to assume any change in the lingual articulation, if it were the case that the inherent glottal gesture for the vowel segment were present and consequently the glottal constriction were narrower for this segment in comparison with the adjacent fricatives. The devoicing of the vowel could take place simply because of some aerodynamic effects under this particular phonetic condition. In the present study, however, this hypothesis is refuted by the following evidence.

In Figure 4, successive frames for the test word /sisee/ containing the devoiced high front vowel [i̥] are presented (series A). For comparison, similar frames for the same test word in which the vowel was voiced are presented in Figure 5 (series B). It is clear in Figure 4 that the vocal cords are set apart throughout the period from the initial [s_j] to the end of intervocalic [s], with no discernible effects of adduction for the vowel in-between. This finding compares with the photoglottographic data reported by Sawashima* who observed that the glottis stayed open for the devoiced vowel as well as for the voiceless consonants.

In contrast to the devoiced cases, the glottis closes and the vocal cords are set into vibration for the voiced period of /i/ in the "voiced cases of [i]" (see Figure 5). Similar results have been obtained for the voiced-devoiced contrast of high vowels in other test words. From these findings, it can be concluded that there are definitely different physical states of the vocal cords for the devoiced and voiced versions of the vowel in the same phonetic environment. It may still, at this point, be argued for that there might be no difference in the motor command for the laryngeal movement between voiced and devoiced /i/, but some physical conditions at the level of the glottis may vary resulting in different "passive" effects. This point will be refuted in a later part of the present paper with evidence from the electromyographic data.

It is noted in both Figures 4 and 5, that the glottis appears to be closed tightly for a certain period of time at the end of the utterance. It is likely

---

* M. Sawashima, "Devoiced Syllables in Japanese - A Preliminary Study of Photoelectric Glottography," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 3, 35-41(1969).

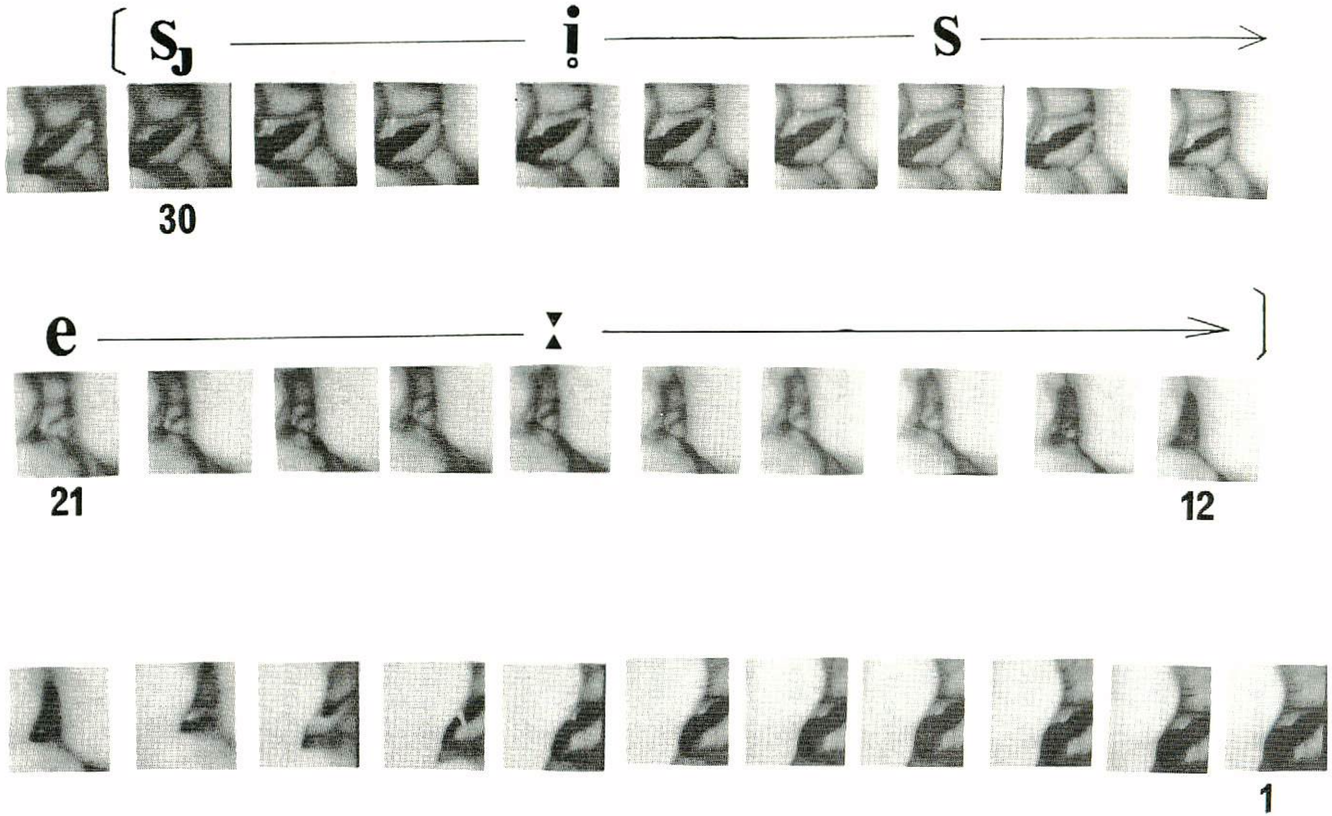


Figure 4. Successive frames for the test word /sisee/ containing the devoiced vowel [i̥].

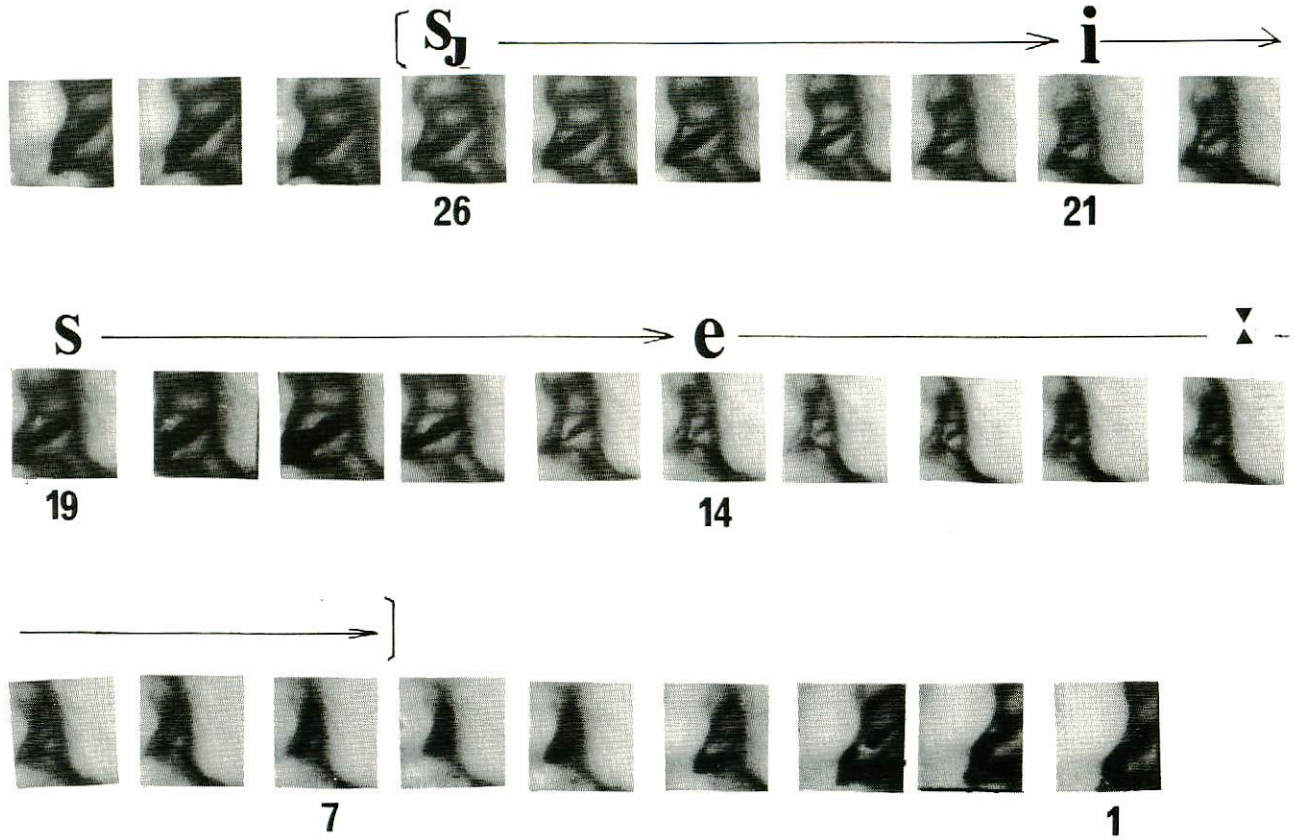


Figure 5. Successive frames for the test word /sisee/ containing the voiced vowel [i].

that the prepared repetition of the test word caused this tight closure of the glottis, preceding the rapid abduction for a short pause between utterances.

Figure 6 shows the raw electromyographic data of the vocalis muscle for the test word /sisee/ uttered with and without devoicing. In the lower sample in which the vowel [i] of the first syllable was uttered without devoicing (series B), a rapid increase in activity of the vocalis muscle is evidenced by an electromyographic burst preceding the onset of the voiced [i].

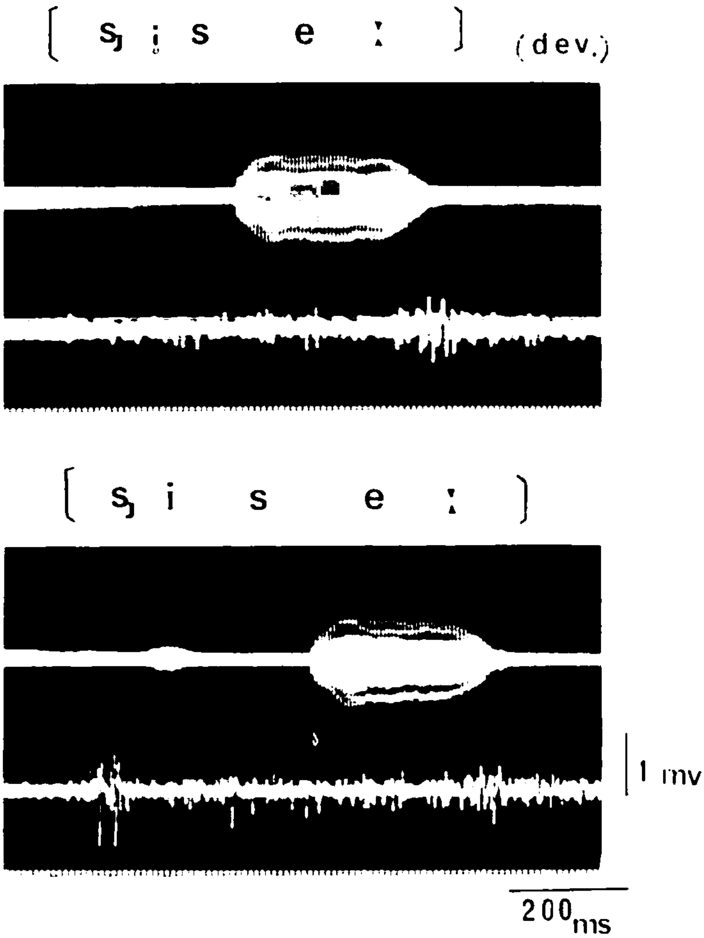


Figure 6. Raw electromyographic data of the vocalis muscle for the test word /sisee/ uttered with (above) and without (below) devoicing. Upper traces: acoustic signal; lower traces: electromyogram.

The activity then decreases and maintains a certain level until it increases again at the end of the utterance. In contrast, in the upper sample in which the vowel [i] is devoiced (series A), there is no burst in the EMG for the first syllable containing the devoiced [i̥]. The difference between the two series can be examined in more detail by appropriately processing the EMG data.

Figure 7 shows the averaged electromyographic data for the same two series (series A and B) together with averaged acoustic signals (the lower traces). The horizontal line in the averaged data indicates the level of muscle activity at complete rest. The thin vertical line in the lower trace indicates the particular acoustic event used as the reference point in the time axis for the averaging process. In this series, the onset of [e] after the intervocalic [s] is taken as the reference point. It is obvious in the averaged data that there is a difference in the pattern of muscle activity between the series with and without devoicing. In the series where vowel [i] in the first syllable is not devoiced (the lower sample of Figure 7:series B), muscle activity begins to increase approximately 120 ms before the onset of voicing for [i] and rapidly forms a peak. Muscle activity then decreases to form a dip and then recovers to a more or less stationary level until it increases again to form another large peak near the moment of the voice cessation.

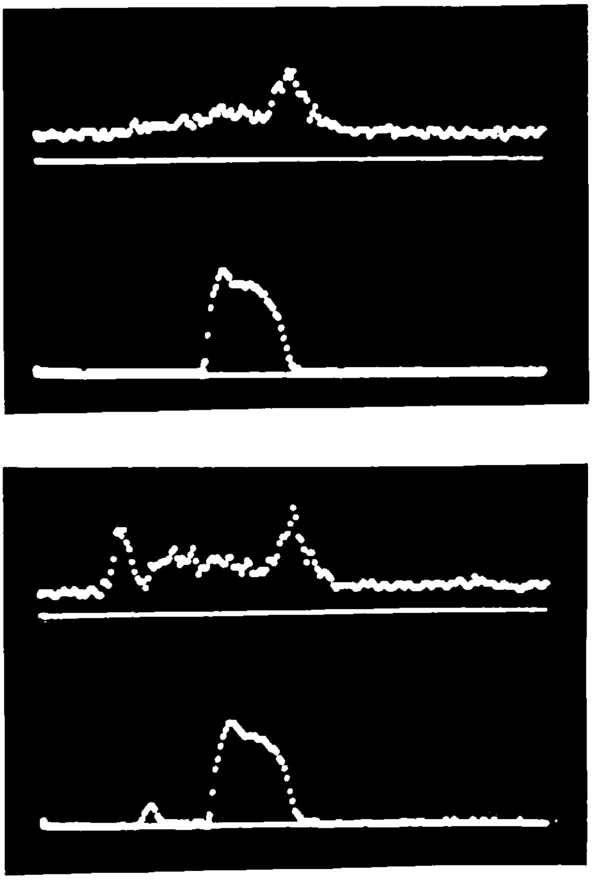


Figure 7. Averaged electromyographic data of the vocalis muscle for the separate series of the utterances of the test word /sisee/ containing the devoiced vowel [i] (above), and the voiced [i] (below), together with averaged acoustic signals. Upper traces: averaged E. M. G.; lower traces: averaged acoustic signal.

Corresponding to the latter peak, a glottal closure accompanying a lowering in pitch is observed by the fiberscope at the end of utterance (Figures 4 and 5). In the series in which the vowel [i] is devoiced (the upper sample of Figure 7: series A), there is no peak in muscle activity for the syllable containing the devoiced vowel. It thus appears that the presence or absence of the phenomenon of vowel devoicing is actually caused by laryngeal motor control. It has to be kept in mind, however, that in the utterances in series A and B, the subject was quite conscious of making the samples consistent through the series in respect to the devoicing factor. This could well have caused some artifacts in the data in this particular point.

In Figure 8, the averaging process was applied to the samples uttered in series C. Samples with and without devoicing occurred randomly in this series and these were separated into two subsets at the time of processing. The characteristics of the difference in the averaged electromyographic patterns between the two subsets are essentially the same as found in the previous series presented in Figure 7. It is clearly revealed in these averaged curves, however, that there are significant activities for the contraction of the vocalis muscle in the earlier phase of the stretch of voicing for

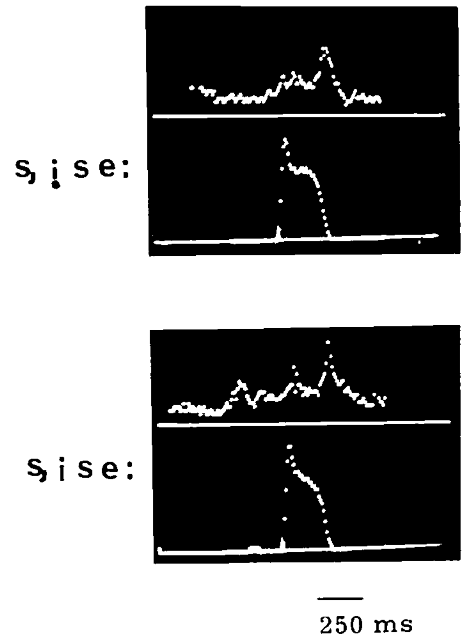


Figure 8. Averaged electromyographic data of the vocalis muscle for the devoiced and voiced utterances. In this series, the subject uttered the test words without paying any attention to the devoicing features. Consequently, samples with and without devoicing occurred randomly. The samples were later separated into two subsets for processing.

the elongated vowel [e:]. It seems that there is an inhibition in the vocalis muscle before the voice cessation, perhaps associated with a slight descent in pitch.

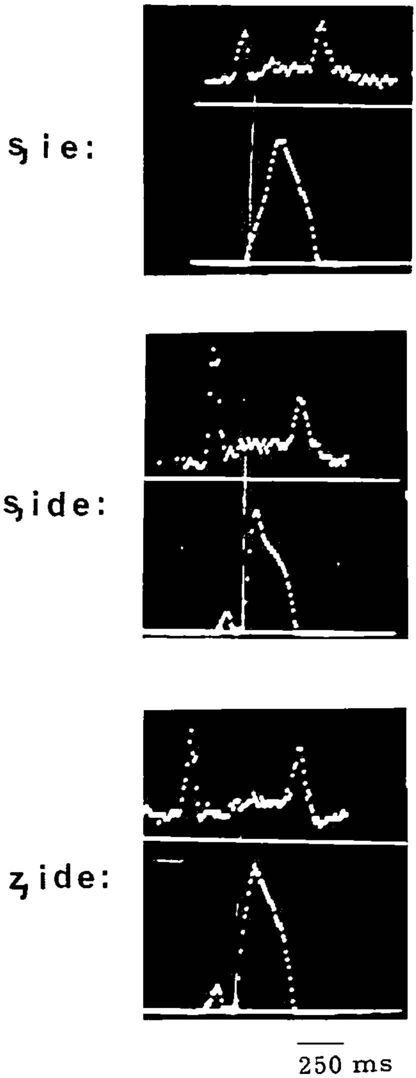
From these electromyographic results, it is now clear that the difference in the laryngeal condition in the presence or absence of devoicing definitely depends on the inherent difference in the motor command, even though the speaker does not make the pertinent distinction.

The devoicing of high vowels under certain phonological conditions in this language may thus be concluded to be a matter concerning the phonological rules. Namely, it seems at this point (perhaps until we learn about other factors that are pertinent to the voicing-devoicing difference) that the choice between the different gestures is made as an optional and free-varying application of a phonological rule.

Figure 9 presents averaged electromyographic data for some other test words. In the case of /si'ee/ and /sidee/, the general pattern of the averaged electromyographic curve essentially resembles that of /sisee/ with voiced [i], and muscle activity decreases once after the initial peak in spite of the fact that voicing continues in these series throughout the utterance. This finding leads us to conclude that the decrease in muscle activity after the initial peak in the case of /sisee/ with the voiced [i], is not due to the presence of a voiceless [s].

In the case of /zidee/, the general pattern resembles that of /si'ee/ and /sidee/, except that the onset of the first peak starts approximately 100 ms earlier than in the other two cases. The activity of the vocalis muscle begins for the voiced [z], in a manner similar to that for a vowel.

Figure 10 illustrates the averaged results of three other test words starting with [z]. The general patterns of these results resemble each other and the muscle activities appear to decrease once after the initial peak, regardless of the difference in conditions for the second syllable.



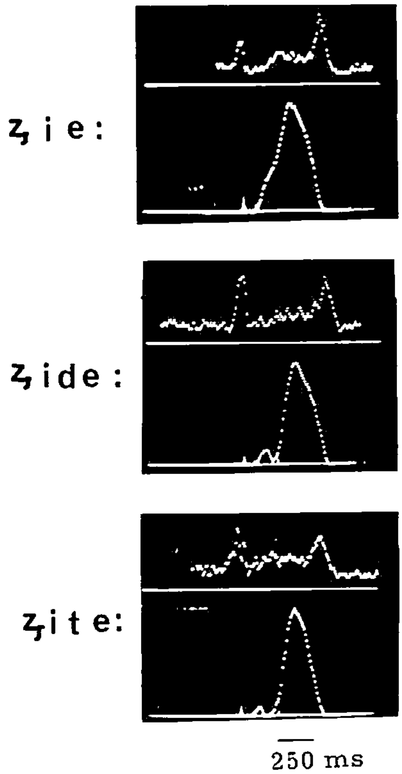


Figure 9 (Left). Averaged electromyographic data on the vocalis muscle for the test words /si'ee/, /sidee/ and /zidee/, together with averaged acoustic signals.

Figure 10 (Right). Examples of averaged electromyographic data of the vocalis muscle for the test words starting with voiced [z].

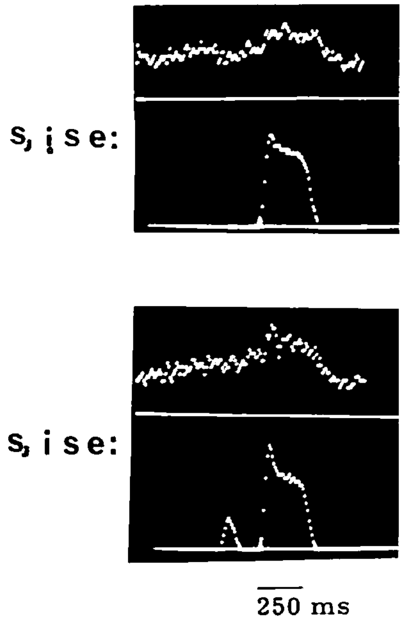
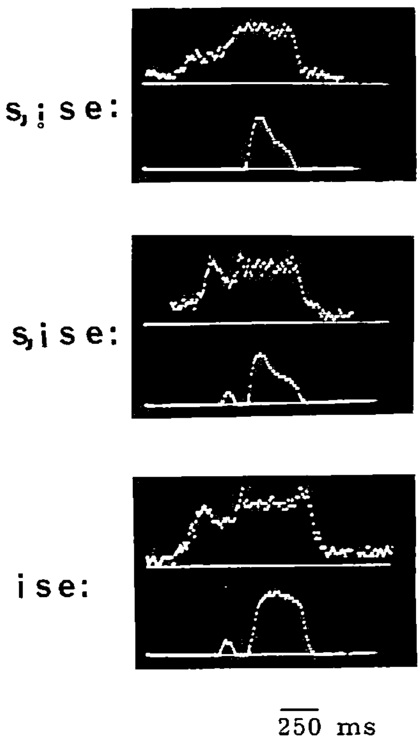


Figure 11. Examples of the averaged electromyographic data for the lateral cricoarytenoid muscle for the test word /sisee/ uttered in the separate series.

Figure 11 shows examples of the averaged electromyographic data for the lateralis muscle in the same utterance as in Figure 7 (in series A and B) for the test word /sisee/ uttered with and without devoicing. The general pattern of electrical activity over time differs considerably from that of the vocalis muscle and the distinction between voicing and devoicing is not clear in these samples.

Averaged electromyographic data for the lateralis muscle of a different subject uttering the samples in series A and B is presented in Figure 12. In these examples, in contrast to the findings for the other subject, the voicing-devoicing distinction is clearly demonstrated in this muscle, too. Namely, there is a peak of activity corresponding to the voiced [i] for the test word /sisee/ uttered without devoicing, whereas the peak is not apparent for the same word with a devoiced [i̥]. These results are comparable to those observed for the vocalis muscle of the first subject, although there is, of course, a considerable difference in the EMG patterns between the vocalis muscle and the lateralis muscle in general.

It is noted that the activity of the lateralis muscle always starts well before the voice onset, somewhat earlier than that of the vocalis muscle, and that, even during the voiceless period of the first syllable of /sisee/ with devoicing of the first vowel, it stays at a certain level of the activity. The activity then stays at a comparatively constant level after it has reached the level for voicing. As the result, the lateralis muscle shows more or less continuous activities throughout the entire period of the utterance, appearing in a trapezoidal shape in the averaged pattern. The activity of the muscle rapidly decreases near the end of utterance. The beginning of the descent appears to correspond to that of the second peak in the vocalis muscle activity which was also observed near the end of utterance.



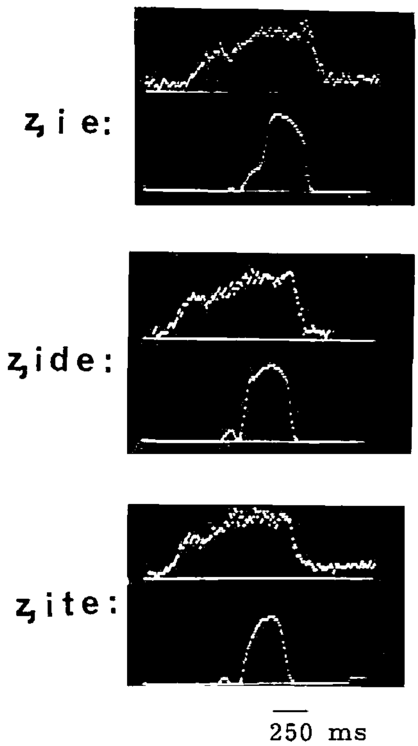


Figure 12 (Left). Averaged electromyographic data for the lateral cricoarytenoid muscle of a second subject. The test words were uttered for the three separate series.

Figure 13 (Right). Examples of the averaged electromyographic data for the lateral cricoarytenoid muscle of the same subject as in Figure 12.

These points are also demonstrated in Figure 13. In these examples, there is a close resemblance in general pattern of the averaged activity regardless of the difference in the phonological content. In other words, the difference in the consonant of the second syllable apparently gives no appreciable difference in the general pattern of the averaged data.

Based on these results, it is considered that the vocalis is particularly active in voicing and has the characteristic of contracting and relaxing in a relatively fast manner, while the lateralis muscle shows relatively continuous activity during the entire period of the utterance. If this assumption is correct, it may be said that the lateralis muscle acts in part (or mainly) for the speech utterance mode as such, while the vocalis muscle plays some specific roles for the initiation of voicing within the utterance.

The results of the present study suggest that there might be functional differentiation in the group of the so-called adductor muscles of the larynx regarding their roles in speech, although interindividual variations in the mode of the activities of the pertinent muscles may also have to be taken into consideration in some cases.