Ann.Bull.RILP(1984)No.18, 153-159

AN INTRAORAL PRESSURE STUDY OF ESOPHAGEAL SPEECH

Hajime Hirose, Kiyoshi Honda and Masayuki Sawashima

In the last issue of this Bulletin, the results of an analysis of the articulatory behavior and its physiological background of skilled esophageal speakers in terms of the voicing distinction were reported by the present authors. As a follow up to that investigation, the present report presents the results of an analysis of the intraoral pressure recorded during the production of a series of test words uttered by selected skilled esophageal speakers.

The measurement of intraoral pressure during speech production in normal subjects has been attempted by many investigators.[1-5] For example, Subtelny, Worth and Sakuda[1] measured intraoral pressure during the production of a variety of speech samples by 30 normal subjects based on the concept that the study of intraoral pressure and airflow was useful in developing a better understanding of the total speech process and consonant production in particular. According to their results, a statistical comparison of the pressure measurement for voiced and voiceless consonants confirmed the generalization that voiceless consonants are produced with a greater amplitude of oral pressure, over a longer period of time, than their respective voiced counterparts.

Lisker[2] measured supraglottal (intraoral) pressure in ten readings by a single subject of a list of twenty short English phrases containing stops in initial, medial and final positions and under various conditions of stress. He found that the voiceless stop group and voiced stop group differed in their mean peak pressure, but, that the difference was not equally significant in all contexts, in that the two categories showed peak pressure overlapping very extensively both initially and medially before a stressed vowel.

Intraoral pressure data obtained from esophageal speakers with reference to the voicing distinction have been very scarce. Isshiki and Tanabe[6] measured intraoral pressure in an electrolarynx speaker and found that intraoral pressure associated with the isolated production of /p/ syllables was very high, reaching as high as 40 - 60 cmH_2O, while that for /b/ syllables was quite variable. They interpreted the high intraoral pressure associated with voiceless stop production as a result of the quick and tense movements of the supraglottal articulators to produce audible plosive sounds. As they claimed, even a minimum volume change in the vocal tract of laryngectomized person may cause greater pressure variation than in normal subjects, since the size of their air-filled cavity is much smaller and less distensible compared to normal subjects having the lung as the organ of air supply.

In the present study, the intraoral pressure of skilled esophageal speakers was recorded during the production of test utterances together with other physiological parameters and acoustic signals to further explore the physiological background of their articulatory behavior.

Procedures

1. Subjects

Three male esophageal speakers served as the subjects of the present study. They ranged in age from 49 to 66 years, and the length of the postoperative period in each case was more than 9 years.

In addition, a normal subject of 50 years of age served as a control subject in which only intraoral pressure was recorded.

2. Recordings of physiological parameters

Intraoral pressure was picked up with a polyethylene tube 15 cm long and of approximately 0.2 cm internal diameter. The tube was inserted through a nostril to the mesopharynx, while the outer end was connected to a differential pressure transducer (Toyota PD-104S). The pressure signals were amplified and recorded on a multichannel PCM data recorder(Model RP 882, NF Electronic Instruments).

Together with the pressure signals, variations in the neoglottal opening were estimated using a photoglottographic system, in which illuminating light was directed through a fiberscope inserted through the other nostril, and the light modulated by the gesture of the neoglottis was sensed by a phototransister attached to the skin surface of the anterior neck above the tracheostoma. The fiberscope was also used to monitor the gesture of the neoglottis during speech production on a video screen.

A pair of surface electrodes were placed on the skin of the anterior neck slightly above the level of the neoglottis for recording EMG signals from underlying muscles. The photoglottographic and EMG signals were recorded on separate channels of the PCM data recorder simultaneously with the pressure and speech signals. The recorded signals were subsequently reproduced and fed to a photocorder (Type 2932, Yokogawa Electric Works, Ltd.) for obtaining graphic traces of the time course of each parameter.

3. Test words

The subjects were required to utter the following Japanese words in the frame "______desu (that is _____)" five to ten times each.

/seHseH/, /seHteH/, /seHzeH/, /deHdeH/, /peHpeH/,
/beHbeH/, /eQpeH/, /siseH/, /jiQseH/, /jiQteH/

Results

Fig. 1 compares the photocorder traces for selected utterance samples of /seHteH/(left side) and /deHdeH/(right side) obtained from one of the three esophageal speakers. The curves are, from the top, EMG, photoglottogram(PGG), intraoral pressure(Po) and audio signal, respectively.

In the case of intervocalic /t/, there is an elevation of intraoral pressure, which is shown here as a downward shift in the Po trace, simultaneously with the opening gestures of the neoglottis, which are also shown as a downward shift in the PGG trace. The time courses of the two traces are very similar, and the Po curve shows a trapezoid contour corresponding to the stop closure period. In the case of intervocalic /d/, on the other hand, the neoglottis does not appear to show an opening gesture, while the intraoral pressure gradually rises to show a triangular contour, the peak value of which is less than that for /t/. These patterns are essentially similar to those obtained in normal subjects. It can be noted in the EMG traces of this particular subject that there is a transient increase in EMG activity immediately before the stop closure for /t/, while no such activity is seen for /d/.

In Fig. 2, the time courses of the three physiological parameters for a selected sample of /seHseH/ are compared with those for /seHzeH/. As for the patterns observed in Fig. 1, the intraoral pressure curve shows a trapezoid time course of pressure elevation during the production of the intervocalic voiceless fricative /s/, corresponding to the glottal opening gesture, while that for voiced /z/ shows a triangular course with almost no opening of the neoglottis. Transient EMG activity was also recorded immediately before the consonantal period for the production of intervocalic /s/ but not for /z/.

In Fig. 3, the values of the peak intraoral pressure obtained from three esophageal speakers are plotted for different types of intervocalic consonants. As a control, the mean value of the peak intraoral pressure obtained from the normal subject is shown together with a range of 2 x SD. It was revealed that the peak intraoral pressure values for the voiceless consonants were significantly higher in the esophageal speakers than in the normal, while the values for the voiced cognates obtained from the esophageal speakers were rather variable and close to those of the normal.

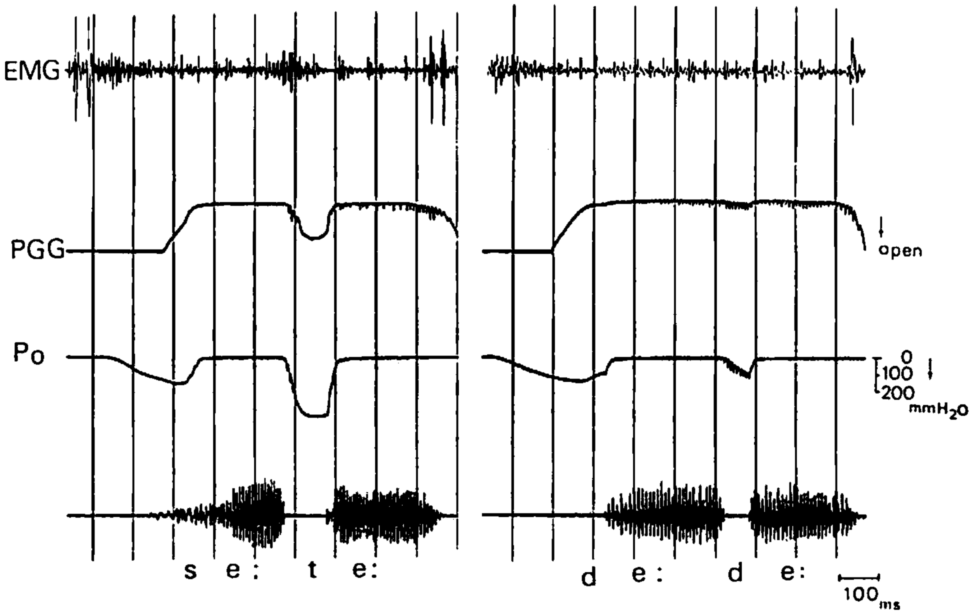

Fig. 1 Simultaneous recordings of EMG, photoglottography(PGG), intraoral pressure (Po) and audio signals for a selected utterance of "/seHteH/ desu" (left) and "/deHdeH/ desu"(right) produced by an esophageal speaker.

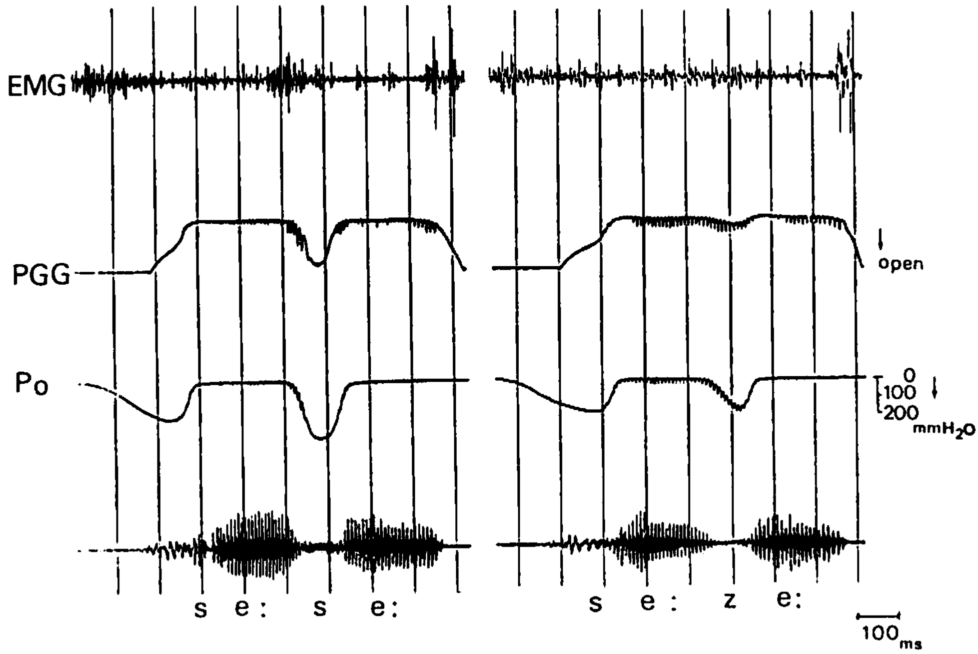

Fig. 2 Simultaneous recordings of EMG, photoglottography (PGG), intraoral pressure (Po) and audio signals for a selected utterance of "/seHseH/ desu" (left) and "/seHzeH/ desu"(right) produced by an esophageal speaker.

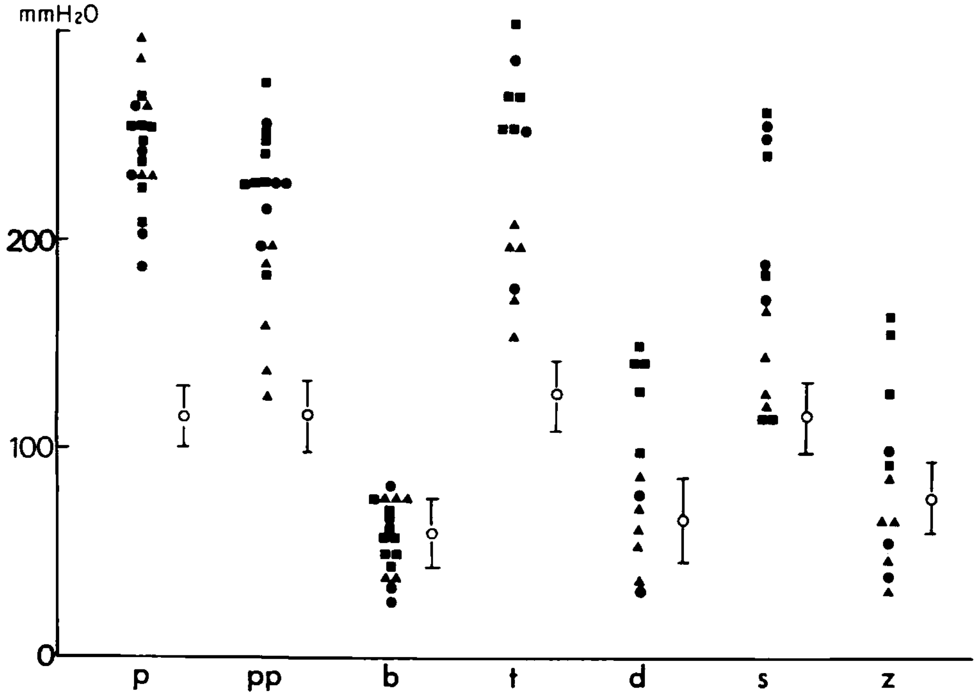

Fig. 3 Peak values of the intraoral pressure (Po) obtained from three esophageal speakers for different types of consonants and a geminate. Each subject is indicated by a filled circle, triangle or square, respectively. The unfilled circles represent the mean of the peak Po obtained from a normal subject, and the vertical bars in-dicate a range of 2 x SD.

Comments

The intraoral pressure difference observed in normal subjects with respect to the voicing distinction has mainly been ascribed to glottal conditions.[6] In the case of voiced consonant production, glottal adduction valves the air from the subglottal space into the supraglottal cavity and, as a result, the transglottal airflow is limited. The relatively slow airflow rate and the short closure duration for voiced consonants do not allow sufficient time for a build-up of intraoral pressure comparable to that characteristic of the voiceless cognates, which allow more transglottal airflow through the open glottis.

In the esophageal speakers examined in the present study, the values of the intraoral pressure for voiceless consonants were significantly higher than those for the normal control, although the general time course of the pressure curves were apparently similar. Since it was confirmed in the present study that the neoglottis is open for voiceless consonant production in esophageal speech, the pressure levels above and below the neoglottis must quickly be equalized during the closure period.

It has been reported that the intraesophageal pressure of laryngectomized persons during phonation is much higher than the subglottal pressure of normal subjects sustaining phonation[7-9]. Thus, the markedly high intraoral pressure of laryngectomized persons during voiceless consonant production can be taken as a result of increasing intraesophageal pressure. However, the fact that high intraoral pressure for voiceless consonants was noted even in an electrolarynx speaker[6] would seem to suggest that supraglottal control can also be responsible for a high intraoral pressure value for voicless consonants.

As mentioned by Diedrich[10], some researchers have held the opinion that air can be injected into the esophagus during the articulation of voiceless consonant sounds. If this were the case, an intraoral pressure increase could overcome the intraesophageal pressure to allow a reverse flow of air into the esophagus. This would imply that the apparent neoglottis opening was a "result" of the intraoral pressure increase rather than its "cause".

However, a careful inspection of the time courses of the glottographic and pressure curves in Figs. 1 and 2 indicates that the neoglottis apparently starts to open at almost the same time as the intraoral pressure increase. In other words, the opening of the neoglottis begins when the intraoral presure is still at a relatively low level.

Further, our preliminary x-ray cinematographic analysis (which is not described in detail in the present paper) revealed that radio-opaque material in the hypopharynx did not appear to be pushed into the esophageal lumen during the period of voiceless consonant closure, though the hypopharyngeal lumen appeared to be expanded to a considerable extent.

Thus, it is not likely that the neoglottis is pushed open by increased intraoral pressure, although the true mechanism of the apparent neoglottal opening during voicless consonant production is still obscure. The timing of the transient EMG activity observed in one of the esophageal speakers with respect to voiceless consonant production may suggest that there is a sort of "active" mechanism for the neoglottal opening in this particular subject, but this interpretation is still tentative. Since the surface EMG electrodes in this case were placed higher than the level of the neoglottis, the recorded activity may not necessarily reflect the control mechanism for the opening of the neoglottis, but may simply relate to the supraglottal control mechanism enhancing voicelessness. A more detailed EMG study using wire electrodes is needed for a clearer understanding of the control of the neoglottis in esophageal speech.

References

1. Subtelny, J.D., Worth, J.H. and Sakuda, M. (1966); Intraoral pressure and rate of flow during speech. Jour. Speech Hear.Res., 9, 498-518.
2. Lisker, L. (1970); Supraglottal air pressure in the production of English stops. Language and Speech, 13, 215-230.
3. Malecot, A. (1966); The effectiveness of intra-oral air-pressure-pulse parameters in distinguishing between stopcognates. Phonetica, 14, 65-81.
4. Arkebauer, H.J., Hixon, T.J. and Hardy, J.C. (1967); Peak intraoral pressures during speech. Jour. Speech Hear. Res., 10, 196-205.
5. Black, J.W. (1950); The pressure component in the production of consonants. Jour. Speech hear. Dis., 15, 207-214.
6. Isshiki, N. and Tanabe, M. (1972); Acoustic and aerodynamic study of a superior electrolarynx speaker. Folia phoniat., 24, 65-76.
7. Dey, F.L. and Kirchner, J.A. (1961); The upper esophageal sphincter after laryngectomy. Laryngoscope, 71, 99-115.
8. Van den Berg, J., Moolenaar-Bul, A.J. and Damste, P.H.(1958); Oesophageal speech. Folia phoniat., 13, 65-84.
9. Frint, von. T. and Pauka, K. (1965); Erfahrungen mit der intra-oesophagealen Druckmessung bei Laryngektomierten. Monatsschr. Ohrenheilk. Laryngol-Rhinol., 99, 284-288.
10. Diedrich, W.M. (1968); The mechanism of esophageal speech. Ann. N.Y. Acad. Sci., 155, 303-317.