MOVEMENTS OF THE LARYNX IN ARTICULATION OF JAPANESE CONSONANTS

Masayuki Sawashima

In order to observe the laryngeal conditions in speech utterances, we devised a special flexible fiberscope which can be inserted through the nasal passage. Photo-electric glottograms have been obtained with simultaneous cinematography and sound recording.

Experimental Techniques

The fiberscope consists of a hard tip, a flexible optical cable, an optical connector, a light source and a camera (Figs. 1, 2). The cable with its hard tip is approximately 6-mm in diameter. The optical cable contains an ligned fiber-optics bundle for the image guide, and a fiberoptics light guide.

Prior to the insertion of the scope, the subject is anesthetized in the nasal fossa and the epipharynx with 4% Xylocain with Epinephrin. Unless the nasal cavity of the subject is too narrow, the insertion of the optical cable is carried out with ease and causes no discomfort or disturbance to the subject in usual pronunciation. When the tip of the scope reaches near the level of the tip of the epiglottis, the larynx is readily visible. The scope can be kept in position or repositioned by manipulating the optical cable at the nostril. Fig. 3 is a lateral radiography which shows the scope in position. Fig. 4 shows still pictures of the glottis of a normal adult male subject in stationary pronunciations of [ʂ] and [i] in falsetto.

A 150-watt incandecent lamp is used as the light source, and a motion picture with synchronizing time marks is taken with a frame rate of 24 per second. During the process of cinematography, a photo-multiplier tube is placed against the pre-tracheal wall of the subject just below the larynx. The light conducted through the light guide passes through the glottis and transcutaneously reaches the photo-multiplier tube. The output signals of the photo-tube are recorded in different levels in three channels of a 7-channel FM tape recorder. The frequency response of the FM tape recorder is flat within ±1 dB from DC to 5000-Hz.

The AC component of the photo-electric signal is recorded also by a dual-track tape recorder simultaneously with the sound and the timing marks. The block diagram shown in Fig. 5 outlines the experimental setup.

When the glottis is open, a linear relationship is empirically ascertained between the output level of the photo-electric signal and the area of the glottis simultaneously obtained from the motion picture. A shift of the output level can be caused by a movement of the optical cable relative to the larynx. By this reason, visual or photographic monitoring of the glottis is indispensable for the correct interpretation of the photo-electric glottogram.

Results

In the first series of experiments, motion pictures of the glottis in pronunciation of CVCV syllables were analyzed for a preliminary study. The symbol C may represent a lax, tense or geminate tense consonant or /h/.

Fig. 6 illustrates sampled frames of the motion picture for /ze˥ze se˥se/ in relation to the sound spectrogram.*

In the selected frames, from left to right, the vocal cords are shown to be: 1) in the abducted (respiratory) position, 2) in the course of adduction, 3) in the adducted position with the glottis slightly open immediately before the voice onset, 4) in vibration for [e] following the initial consonant [dz] of the word /ze˥ze/ (the blurred edges of the vocal cords indicate the vibration), 5) set apart for the word-initial [s] of /se˥se/, 6) again in vibration for the interconsonantal vowel, 7) again set apart for the intervocalic [s] (cf. the word initial [s]).

Successive frames of the glottis for selected Japanese consonants are shown in Figs. 7 - 12.

In /pe˥pe/ (Fig. 7), the vocal cords are set apart for the closure period of the voiceless stop [p]. There is a drop of the fundamental frequency after the first syllable, and the larynx appears to sink and the glottis looks shortened (also covered by the epiglottis) for the second vowel.

In /se˥se/ (Fig. 8), the vocal cords are in a clearly abducted position at the

* The mark ˥ in the phonemic transcription denotes the position of the "accent kernel", which is manifested in the speech signal as a drop of the fundamental frequency.

onset of the utterance initial [s] , while in the intervocalic [s] they are significantly less abducted.

At the onset of the intervocalic /h/ in /he⌉he/ (the right most frame of the second row in Fig. 9), the vocal cords continue to vibrate while they are set apart. The voicing is clear also in the sound spectrogram of the utterance.

Fig. 10 is a comparison of the intervocalic /h/ and /s/. In /h/, the vibration of the vocal cords continues throughout the non-vocalic period even though they appear to be set further apart than in /s/.

In /be⌉be/ (Fig. 11), the vocal cords are in the adducted position and are vibrating for the closure periods of [b].

In /ze⌉ze/ (Fig. 12), adducted vocal cords leaving a slightly open glottis are observed for some time interval preceding the voice onset for the consonant in the utterance initial position (cf. also Fig. 6). Vocal cords are vibrating for the constrictive period of [z].

In the second series of experiments, the opening area of the glottis was photo-electrically recorded for various consonantal articulations, with the simultaneous 16-mm cinematography. In this case, CVCV syllables were embedded in the carrier sentence "Kore wa ii ____ desu (This is a good ____)," and the motion pictures of the glottis were used for ascertaining the appropriate positioning.

Fig. 13 is a recorded sample continuing /keQke/, where /Qk/ is manifested phonetically as a geminate consonant [kk]. A sound spectrogram, sampled frames of the motion picture and a photo-electric glottogram are shown. An upward shift of the curve in the glottogram represents an opening process of the glottis, and a downward shift a closing process. Figs. 14 - 15 are photo-electric glottograms for various Japanese consonants.

In the constrictive phase of voiceless consonants, the vocal cords are kept apart, and the size of the maximum glottal area generally seems to be greater in fricative consonants than in stops in the same phonological environment. The maximum area for the same phoneme seems to vary appreciably depending on the environment, and also on the speed of utterance.

In the case of geminate voiceless fricatives, in particular, the glottal area is apparently larger compared with the non-geminate counterpart.

The vibration of the vocal cords ceases gradually as the opening movement of the vocal cords proceeds after a closure or constriction of the vocal tract has been initiated. This point is particularly marked in the second syllable of CVCV.

An abrupt buildup of the vocal cord vibration is observed for a vowel portion following a voiceless consonant.

In the intervocalic /h/, the vocal cords are apt to vibrate throughout the consonantal period even while they are widely apart.

In the voiced stops and fricatives, compared to vowel segments, a reduction in amplitude of the vocal cord vibration is observed (Fig. 15). The extend of the reduction apparently depends on the degree of the vowel fract constriction, or presumably the amount of air flow. Thus, no marked change in the glottal condition is observed for nasal consonants.

References

Halle, M. and Stevens, K. N.: "On the Mechanism of Glottal Vibration for Vowels and Consonants," _Quarterly Progress Report,_ Research Laboratory of Electronics, Massachusetts Institute of Technology, No.85, 267-271 (1967).

Malecot, A. and Peebles, K.: "An Optical Device for Recording Glottal Adduction-Abduction during Normal Speech," _Zeitschrift fur Phonetik Sprachwissenschaft und Kommunikationsforschung_ _18_, Ht. 6, 545-550 (1965).

Ohala, J.: "A New Photo-Electric Glottograph," _Working Papers in Phonetics_ _4_, 40-52, University of California, Los Angeles, (1966).

Sawashima, M. and Hirose, H.: "New Laryngoscopic Technique by Use of Fiber Optics," _J. Acoust. Soc. Amer._ 43, 168-169 (1968).

Sawashima, M., Hirose, H., and Fujimura, O.: "Observation of the Larynx by a Fiberscope Inserted through the Nose," _J. Acoust. Soc. Amer._ _42_, 1208 (1967).

Sonesson, B.: "On the Anatomy and Vibratory Pattern of the Human Vocal Folds," _Acta Otolaryng._ Suppl. 156 (1960).

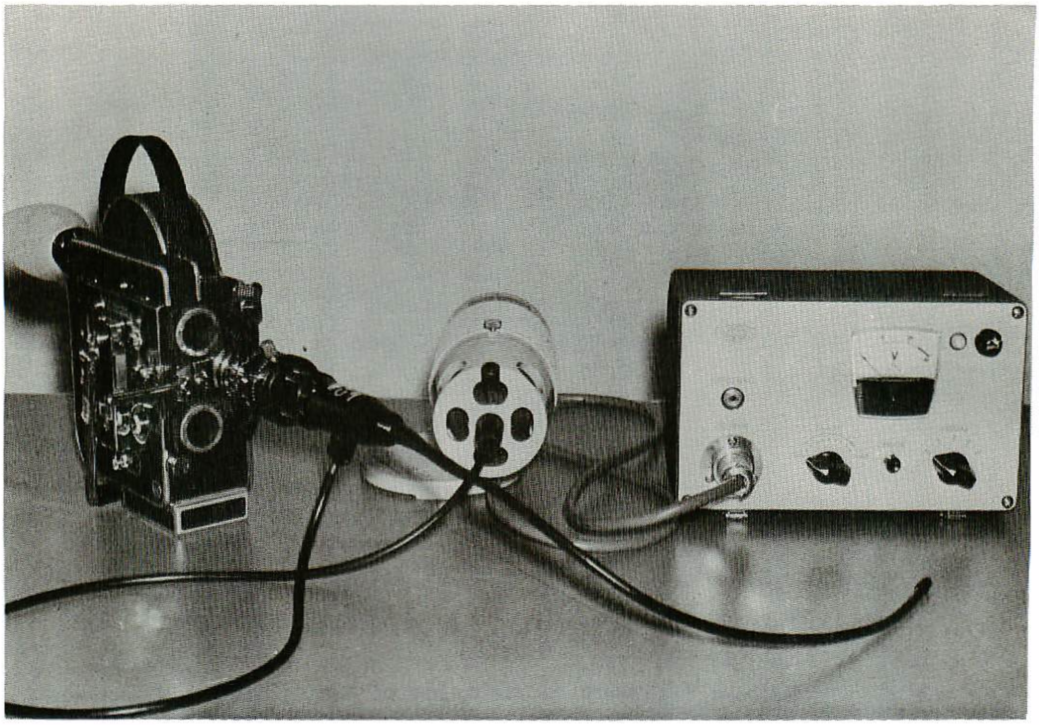


Fig. 1 - A photograph of the fiberscope
with a 16-mm movie camera.

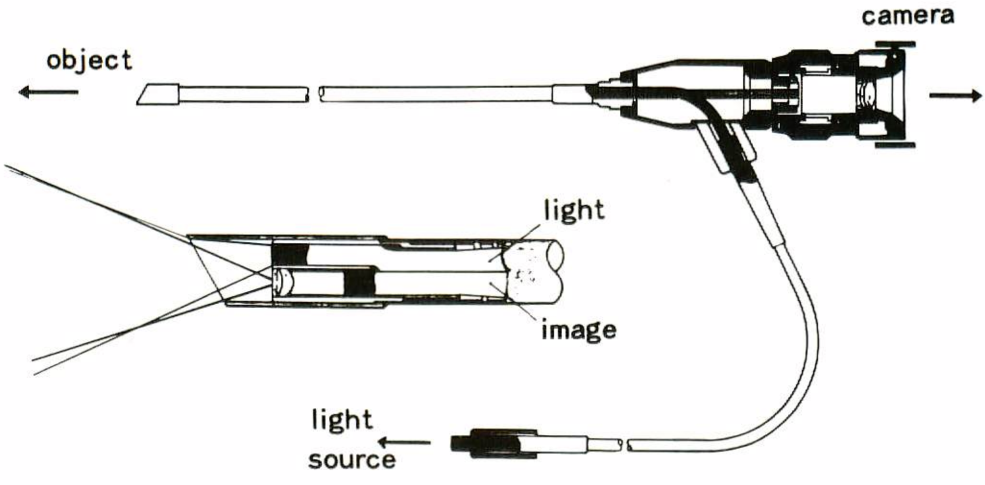


Fig. 2 - Schematic drawing of the structure
of the fiberscope. The hard tip is
magnified and shown in the center
of the figure.

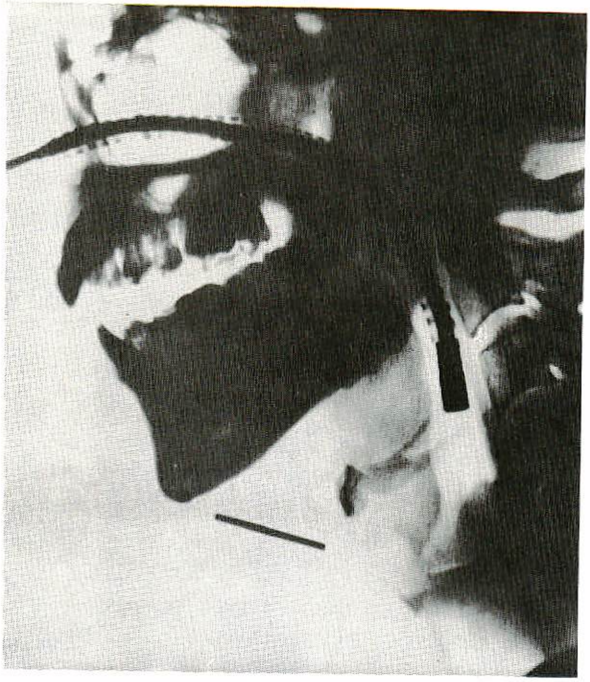


Fig. 3 - Typical positioning of the optical cable (a lateral x-ray picture).

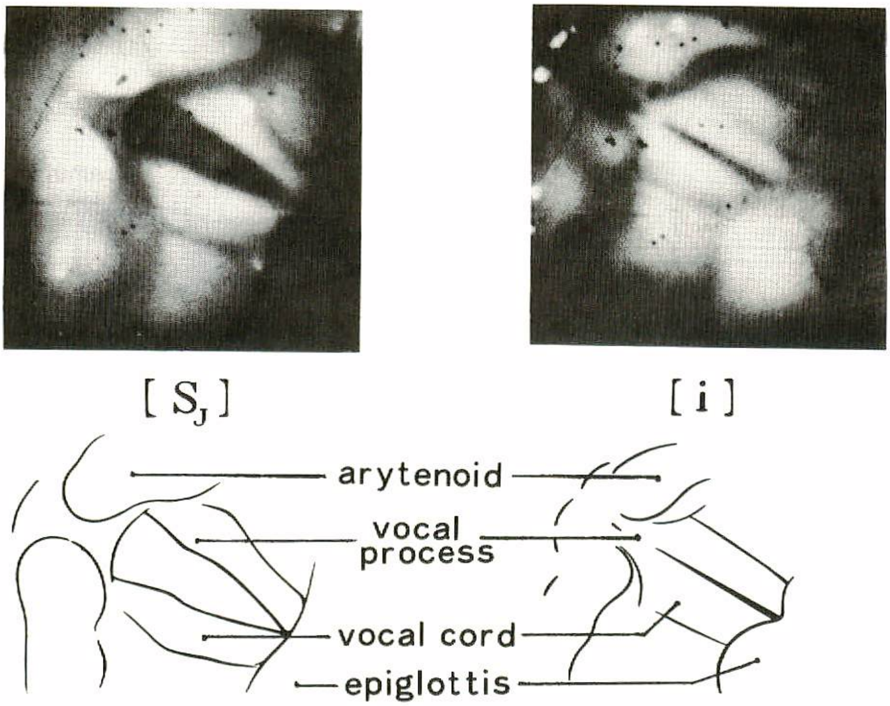


Fig. 4 - Two examples of the larynx viewed through the fiberscope (still pictures).

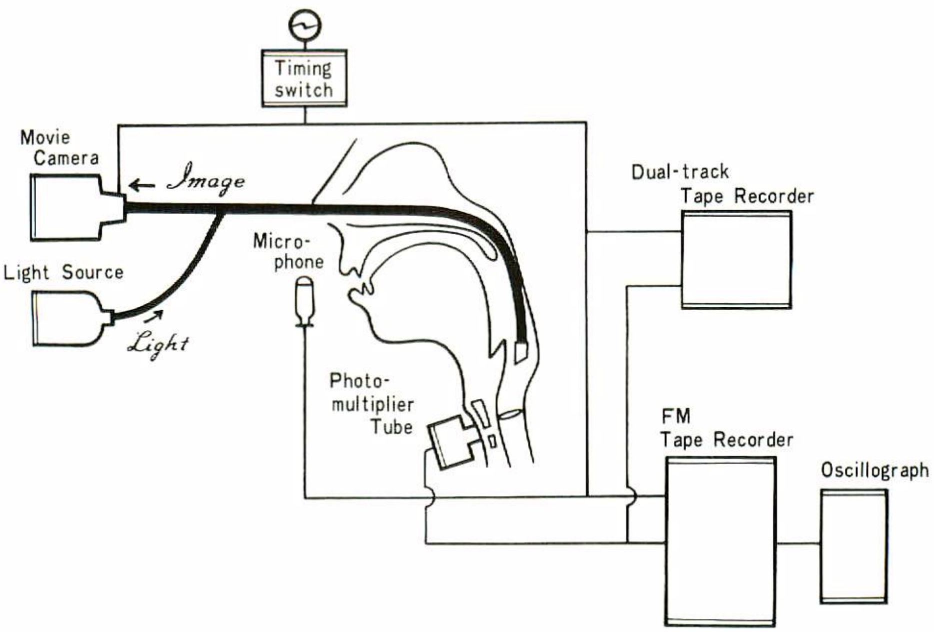


Fig. 5 - The experimental setup for simultaneous use of a movie camera and the photo-electric glottograph

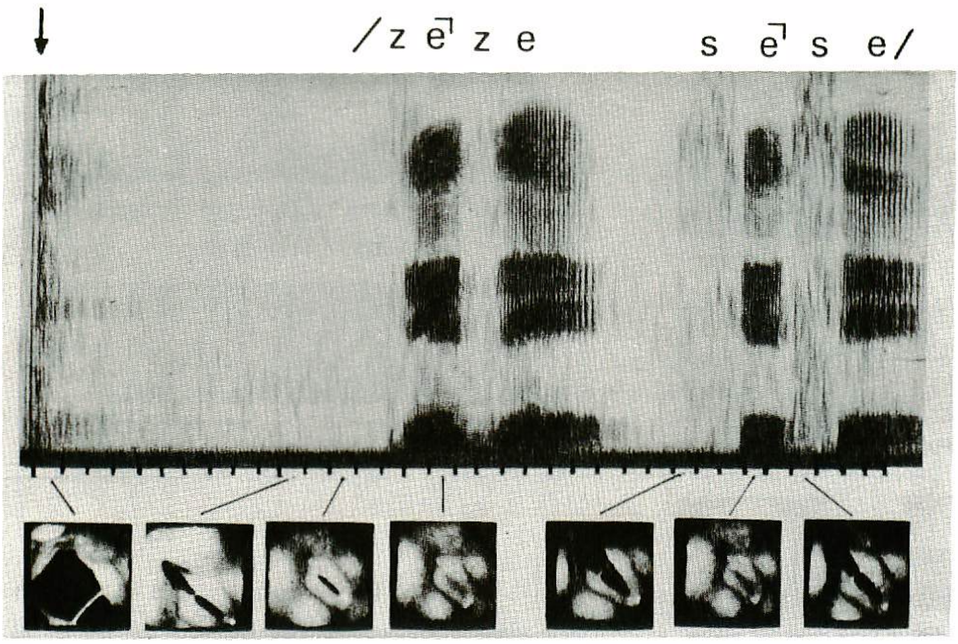


Fig. 6 - Sound spectrogram and selected frames of the motion picture shwoing the movement of the larynx in a pronunciation of /ze ze se se/ (Subject S). Short vertical bars under the spectrogram demarcate time intervals for successive frames of the motion picture. The arrow at the upper left corner indicates the time mark that is placed at the start of utterance for synchronization of sound and frames.

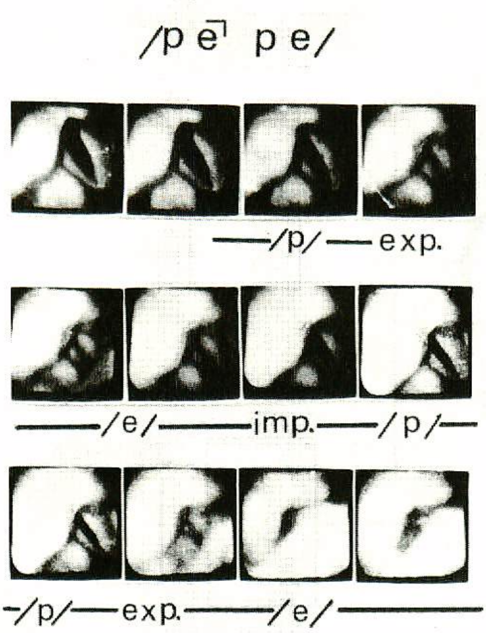


Fig. 7 - Successive frames for /pe˥pe/ (subject F.).

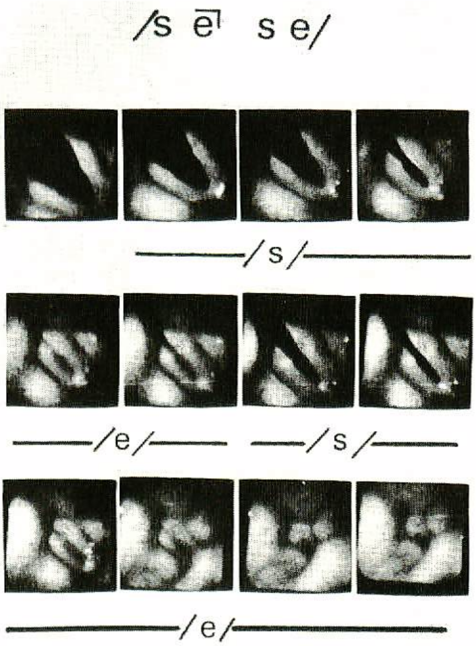
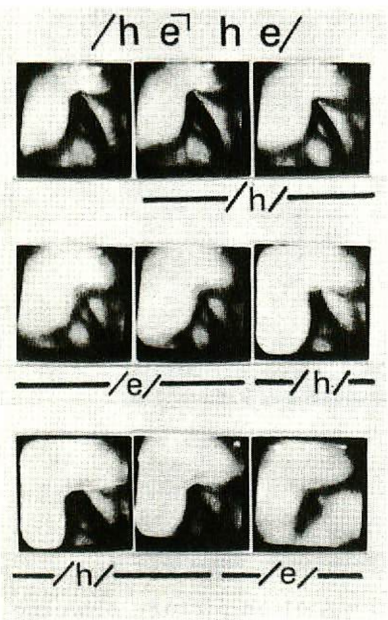


Fig. 8 - Successive frames for /se˥se/ (subject S.).



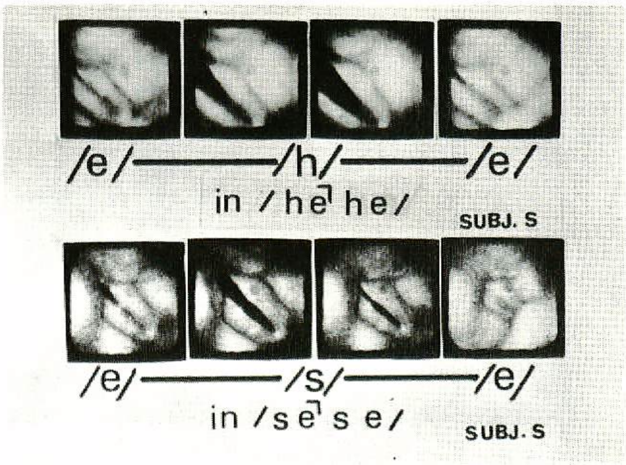


Fig. 10 - Comparison of successive frames for the intervocalic /h/ and /s/ (subject S.).

Fig. 9 - Successive frames for /he˥he/ (subject F.).

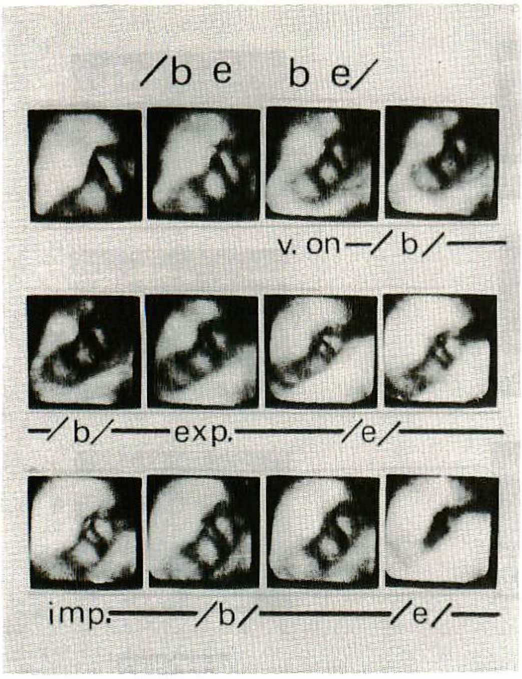


Fig. 11 - Successive frames for /be˥be/ (subject F.)

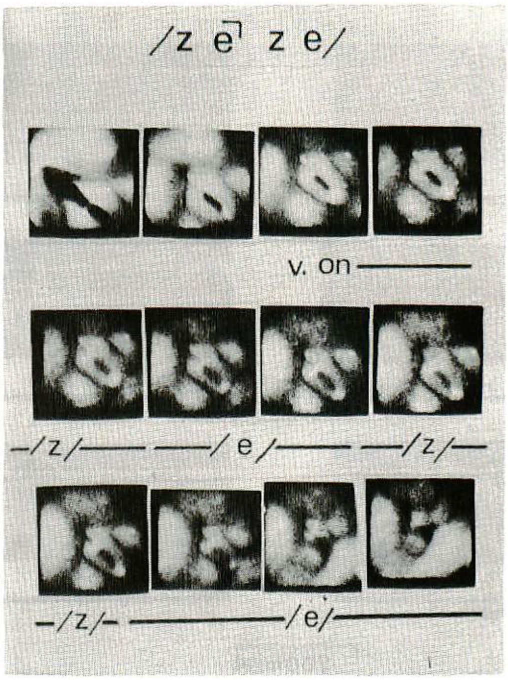


Fig. 12 - Successive frames for /ze˥ze/ (subject S.).

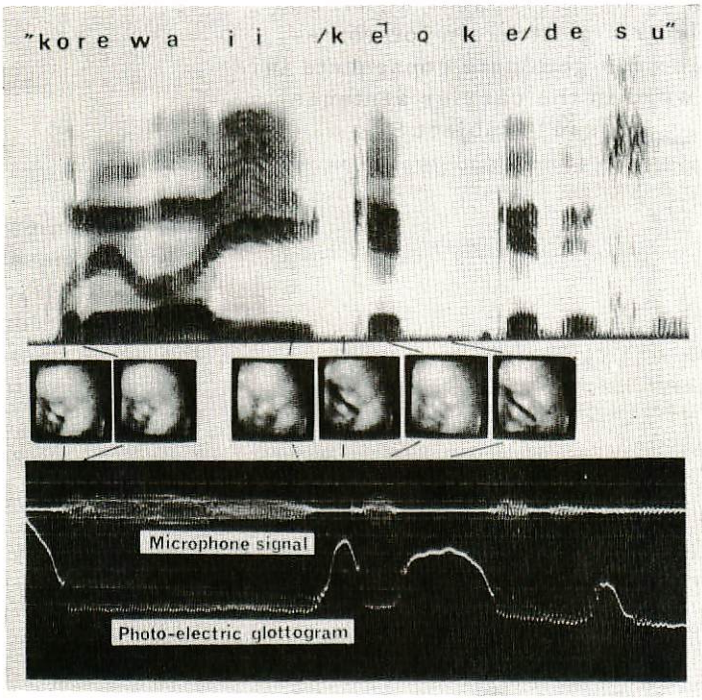


Fig. 13 - A sound spectrogram and selected frames of the motion picture and a photo-electric glottogram taken simultaneously for an utterance of a sample sentence (subject S.).

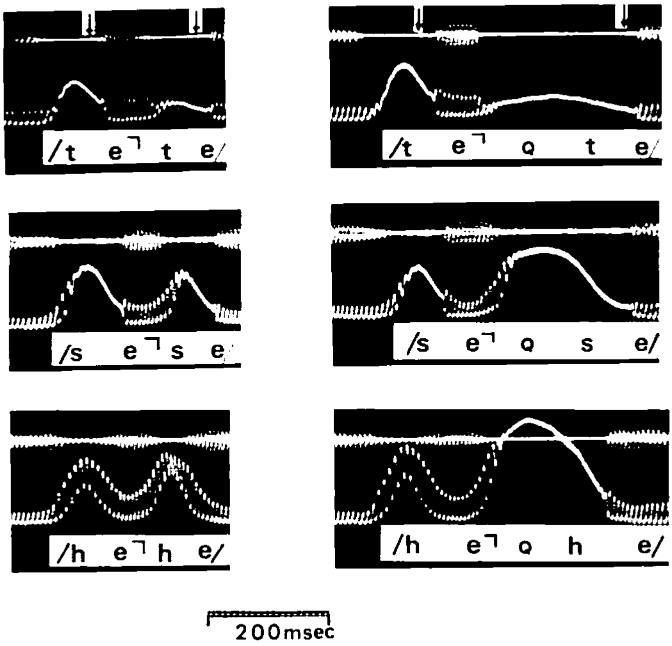


Fig. 14 - Comparison of the microphone signal and the photo-electric glottogram for voiceless and voiceless geminate consonants in a /CV CV/ word in the carrier sentence "kore wa ii ____ desu" (subject S.).

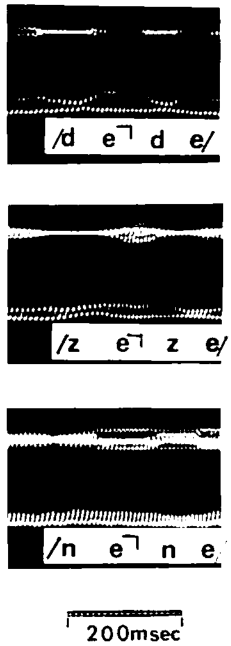


Fig. 15 - Comparison of the microphone signal and the photo-electric glottogram for voiced consonants in the same phonological condition as in Fig. 14 (subject S.).