Ann.Bull.RILP(1989)No.23, 1-9

Vocal Cord Vibration at the Release of the Consonant -Observation by the High-speed Digital Image Recording-

Shigeru Kiritani, Hiroshi Imagawa and Hajime Hirose

In order to facilitate high-speed recording of vocal cord vibration, we have developed a system of digital image recording consisting of an image sensor and a digital image memory. By using a solid type endoscope, recordings of vocal cord vibration during sutained phonation have been made. Furthermore, recording of vocal cord vibration during running speech was made possible by combining the system with a fiberscope1)-3). In order to fully expolit the advantage of the fiberscopic system, it is desired to have a digital image memory of large capacity so that continuous recording of vocal cord image during an utterance can be performed.

In the present study, a special large size image memory was constructed which enables data recording for an utterance of up to one second. A few results of observation of vocal cord vibration during the release of intervocalic consonants will be reported.

Image recording system

Figure 1 shows a block diagram of the image recording system. The fiberscope is connected to a single lense reflex camera. Inside the camera body is located a CCD sensor which records the image from the fiberscope. The video signal from the CCD sensor is sampled by a high-speed A/D converter and stored in a digital image memory. In order to realize high-speed image recording, only 128x16 picture elements in the CCD sensor are sampled although the sensor contains 350x250 picture elements. By using an especially designed fiberscope which is equipped with a larger number of light guide fibers than a conventional scope, and the 300W Xenon lamp, image recording at a rate of 2000 frames per second was achieved.

In the present system, a special image memory of 8MB was constructed which functions as a two port memory attached to a personal computer. During image recording, video data from the A/D converter are written into the image memory. After data recording, the CPU accesses the image memory and transfers frame by frame the image data to a video RAM on the frame memory board. This board is also attached to the personal computer and generates ordinary video signal of the image data in the video RAM. Personal computer can rewrite the data in the video RAM at a rate of 10 frames per second. Thus, a slow motion display of 20 times the real speed is produced on a video monitor. Several different mode of image display and analysis are possible in the present system.

Results

Figure 2 shows vocal cord vibrations during the production of an intervocalic consonant /p/ (word initial /p/ in the utterance /i: pepe desu/). During the period from frame No.246 to frame No.300, which corresponds to the implosion of /p/, there are three vibratory cycles which have a clear closed phase. The closed phase becomes shorterin each successive cycles indicating that the voval cords are getting seperated for the production of voiceless consonant. Following these vibrations, there can be seen at least two vibratory cycles which do not have the complete closure of the glottis. Moments of maximum glottal narrowing for these two cycles are seen at frame No.300 and No.314.

From frame No.318 to frame No.370, the glottis remains open and there is no appreciable change in the glottal opening. At about frame No.370, narrowing of the glottis starts and the vocal cords start to vibrate. In the first cycle of vibration, glottal closure is incomplete. The moment of minimum glottal opening is observed in frame No.380. In the next cycle, closure of the glottis is complete and in the following vibratory cycles, duration of the closed phase becomes longer.

Figure 3 shows the speech waveform during the implosion and release of consonant /p/, together with the timing of the closed phase of vocal cord vibration. It is noted that in the release period, start of the clear pattern of excitation in the speech wave corresponds to the appearance of the closed phase in vocal cord vibration. The first vibration after the silent period of /p/ does not have a closed phase and does not generate appreciable excitation of the speech wave. Contrary to this, during the implosion of consonant, the excitation pattern of the speech wave decays even when vocal cord vibrations with complete closure of the glottis still continue. It appears that the decay of the exitation pattern in this period is due to the formation of the closure in the vocal tract.

Figure 4 shows vocal cord vibrations during the release of /s/ (word initial /s/ in the utterance /i: sese desu/). At around frame No.380, vocal cords begin to approximate and the first cycle of vocal cord viration starts. In this cycle, the glottal closure is incomplete. As in the case of the post-consonantal voicing for /p/, the second vibration shows complete closure of the glottis (frame No.400) which corresponds to the strong excitation in the speech wave.

Figure 5 shows the speech waveform and the timing of the closed phase of vocal cord vibration for five tokens of /p/ and three tokens of /s/. It can be seen in the figure that the start of strong excition in the speech wave coincides with the start of vocal cord vibration with glottal closure. In the start of post-consonantal voicing, the first cycle of vocal cord vibration generally does not have a closed phase; that is, glottal closure is incomplete. This vocal cord vibration does not produce the strong excitation of a speech wave. However, in the second cycle, complete glottal closure is always attained.

It also can be seen in the figure that, during the following two or three cycles, the duration of the closed phase becomes longer. It should be noted that, due to this change in the duration of the closed phase, the interval between successive closing points of vocal cord vibration is shorter during consonant release than during the period of following vowel. It has been generally noted that in the release period of a voiceless consonant, pitch frequency is higher than in the following vowel. Pitch period is generally determinded with reference to the main peaks in the speech wave which correspond to the excitations at the closing points of vocal cord vibration. Thus, the above mentioned phenomenon (change in the duration of the closed phase in the relase period) can explain, at least in part, the higher pitch frequency in the post-consonantal period.

Figure 6 shows vocal cord vibration for the production of /b/ (word medial /b/ in /i: bebe desu/). Vocal cord vibration during consonant closure and during the preceding vowel are compared. It can be seen in the figure that there is no appreciable difference in the pattern of vocal cord vibration during vowel /e/ and consonant /b/. In the present images, it is difficult to find any apparent indication that the glottal constriction is looser for /b/ than for the vowel. It is noted that that the closed phase is even longer for /b/.

Summary

1) The first cycle of vocal cord vibration in the start of post-consonantal voicing generally does not have complete closure of the glottis. In the second cycle, complete closure is attained which corresponds to the start of strong excitation in the speech wave.
2) Change in the duration of the closed phase during the first few vibratory cycles contributes to a higher pitch frequency in the release period of a voiceless consonant.
3) Vocal conrd vibration during the closure period of /b/ maintains a closed phase. Its closure period is longer than that during a vowel.

References

1) Honda, K., S. Kiritani, H. Imagawa and H. Hirose: High-speed digital image recording of vocal fold vibrations using a solid-state image sensor. Laryngeal function in phonation and respiration (Eds. Baer, T., C. Sasaki and K. S. Harris), College Hill Press, 485-491, 1987.
2) Imagawa, H., S. Kiritani and H. Hirose: High-speed digital image recording system for observing vocal fold vibration using an image sensor, Jap. J. Medical Electronics and Biological Engineering, 25, 284-290, 1987.

3) Kiritani, S., K. Honda, H. Imagawa and H. Hirose: Simultaneous high-speed digital recording of vocal fold vibration and speech signal, Proceedings ICASSP, 1633-1636, 1986.

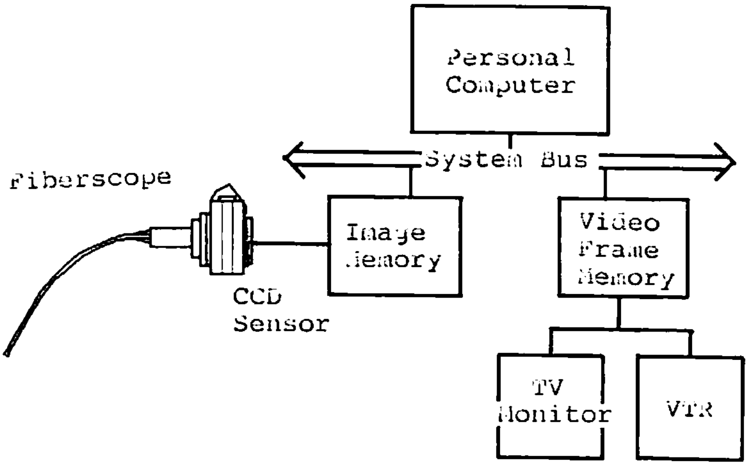
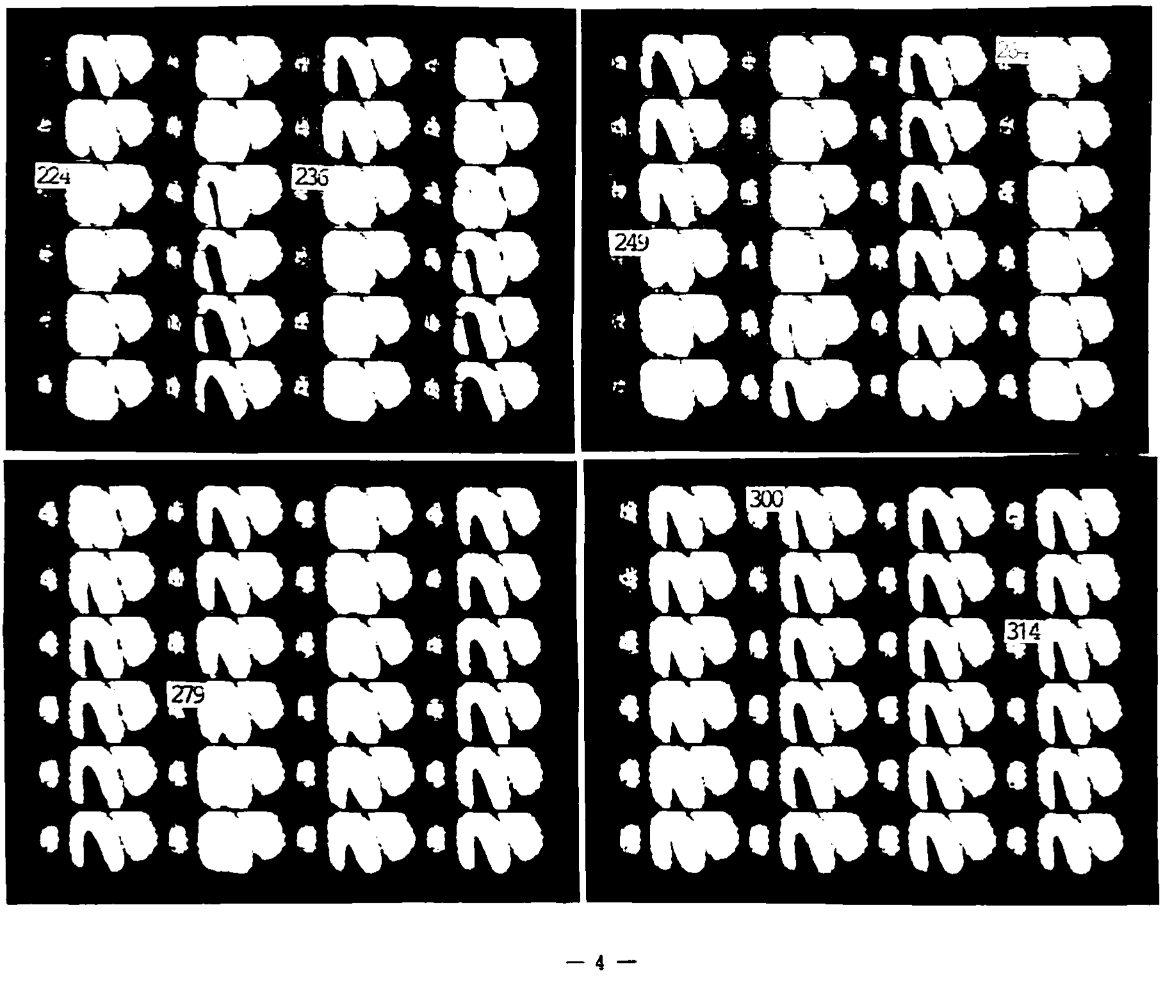


Figure 1 Block diagram of image recording system.



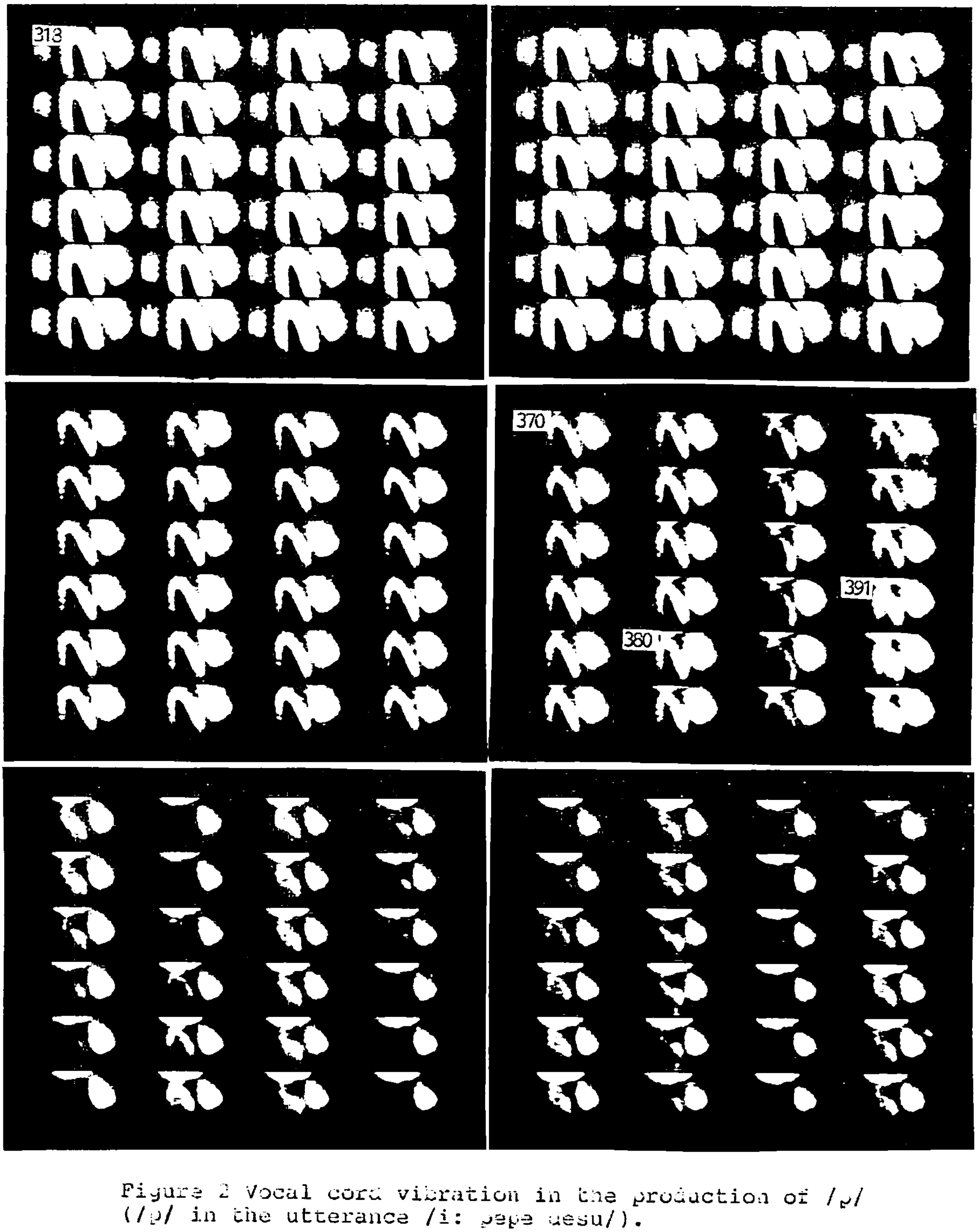


Figure 2 Vocal cord vibration in the production of /p/ (/p/ in the utterance /i: pepe desu/).

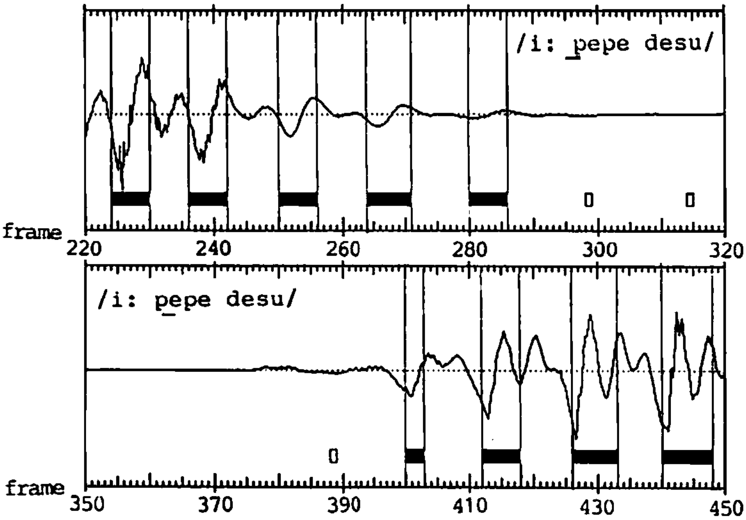


Figure 3 Speech waveform and closed phase of vocal cord vibration in the onset and release of /p/.

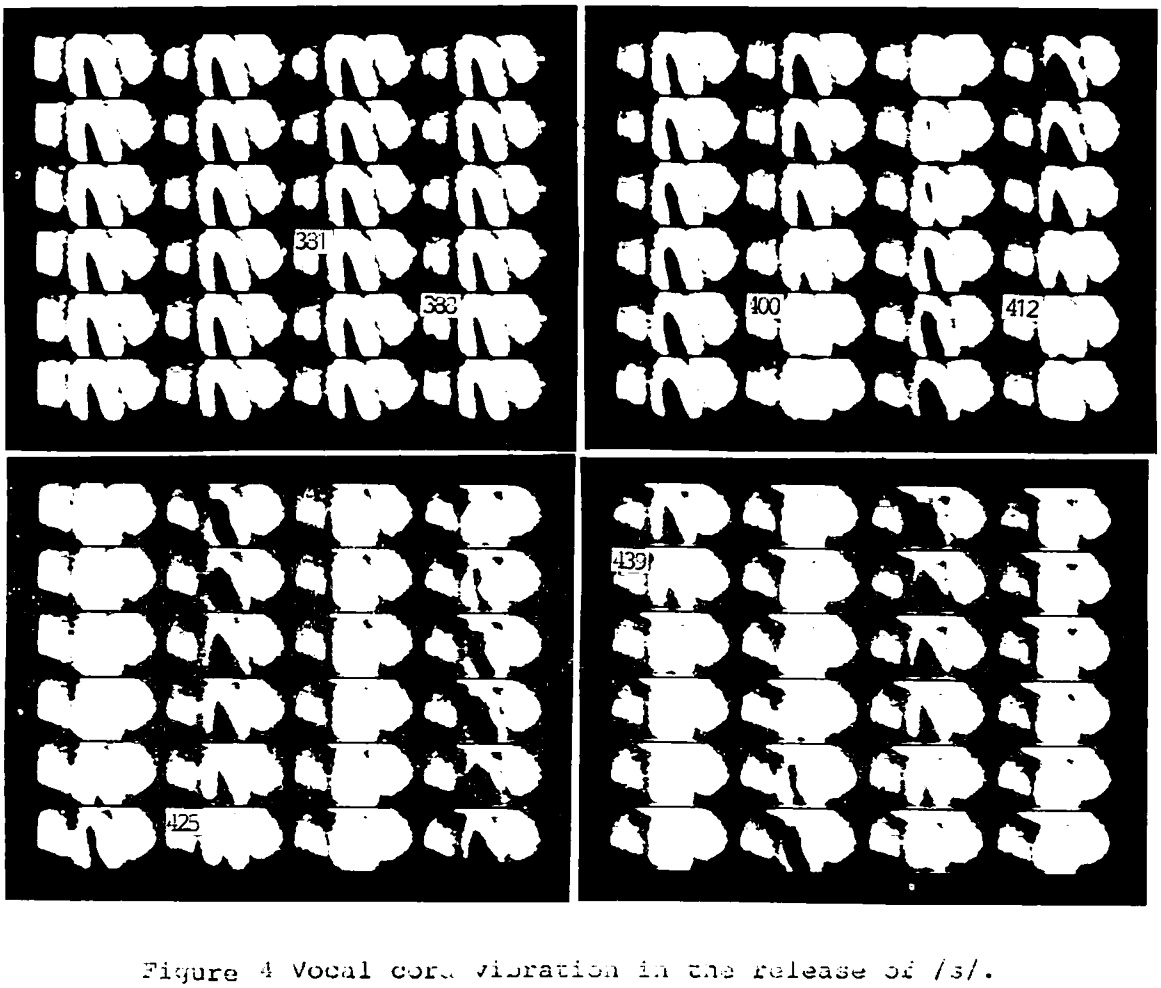
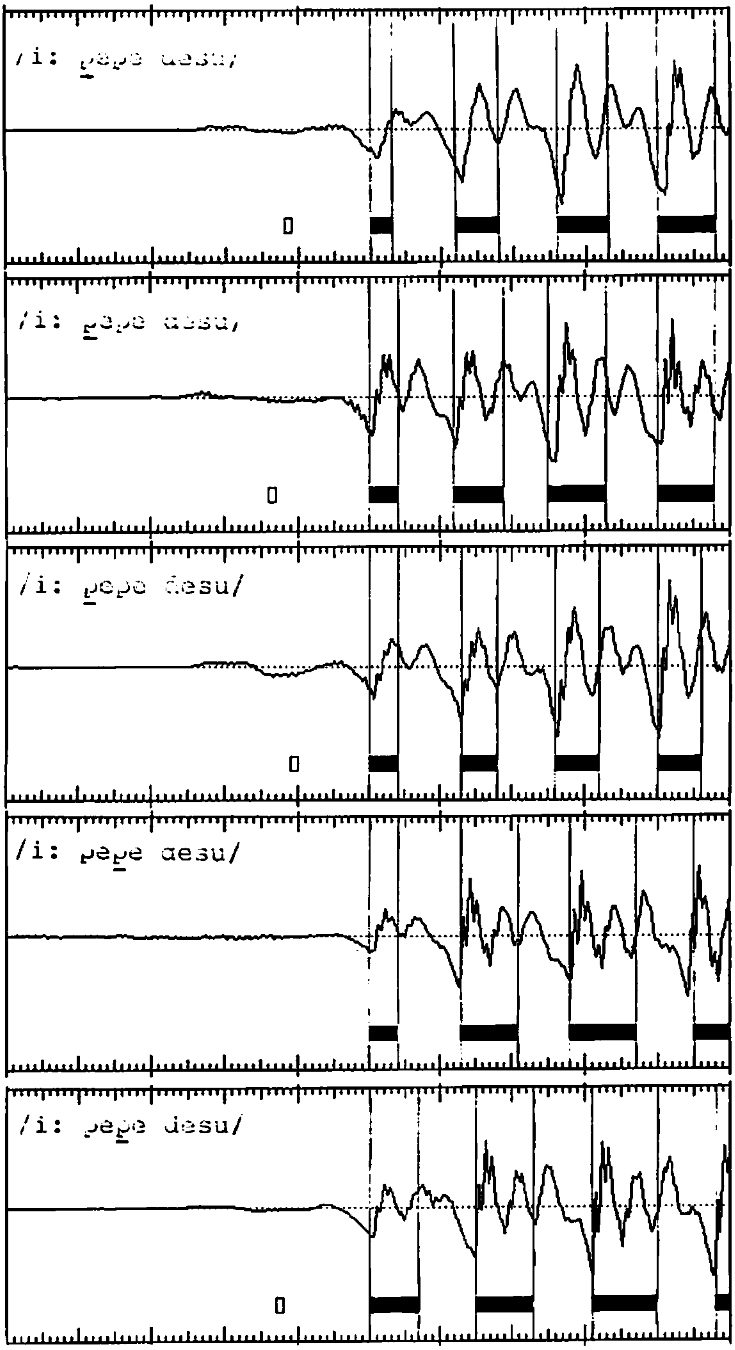


Figure 4 Vocal cord vibration in the release of /s/.


/i: pepe desu/
/i: pepe desu/
/i: pepe desu/
/i: pepe desu/
/i: pepe desu/

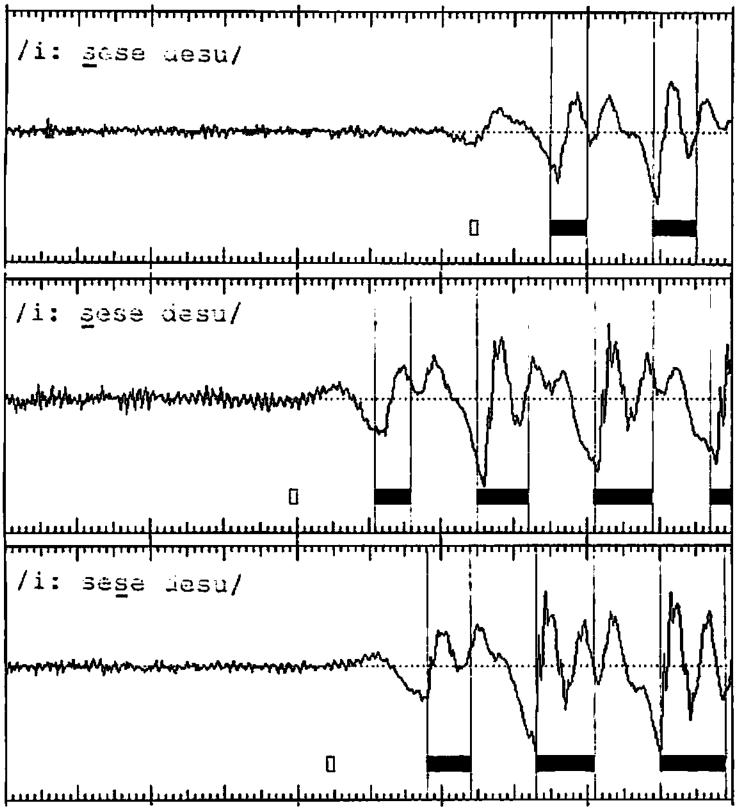
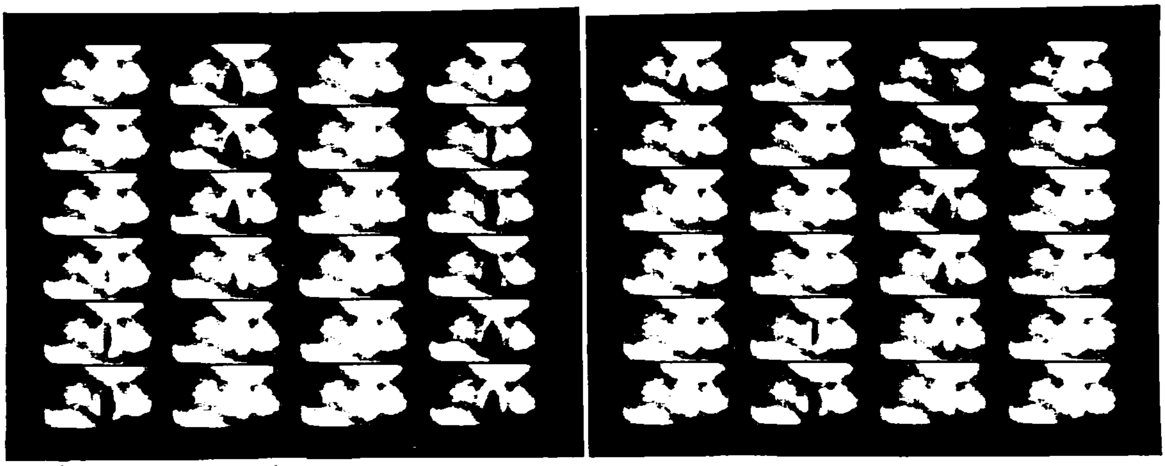


Figure 5 Speech waveform and closed phase of vocal cord vibration at the release of /p/ and /s/.

/i: bebe desu/



/i: bebe desu/

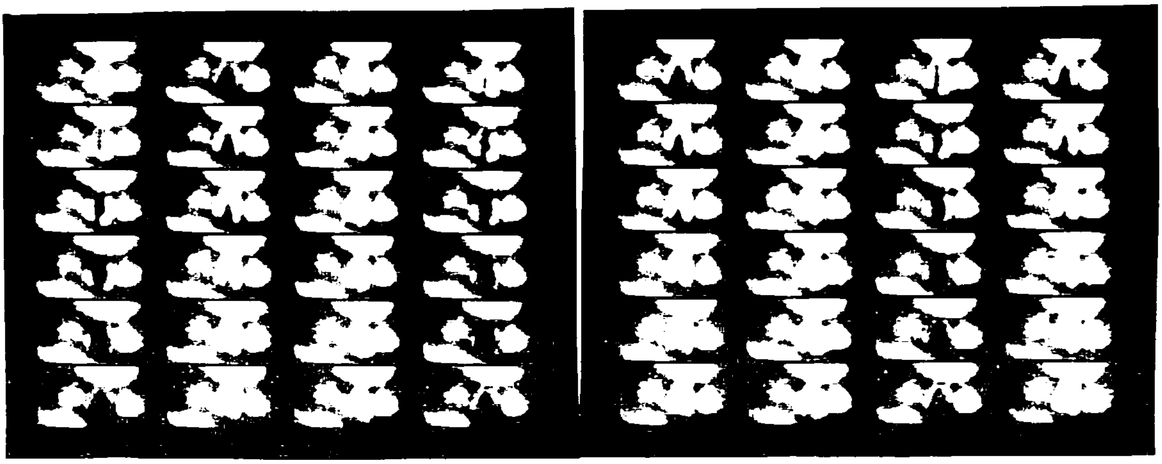


Figure 6 Vocal cord vibration in the production of /i: bebe desu/. Voel and closure period of /b/.