High-speed Digital Imaging of Vocal Cord Vibration at Voice Onset in Consonants

S. Kiritani, H. Imagawa and H. Hirose

Introduction

The high-speed digital image recording system using a fiberscope developed by the present authors is very valuable in that it makes possible observations of vocal cord vibrations for consonants. However, to fully exploit the advantage of this system, it is essential to have a large image memory for data storage so that the system can record data from continuous speech. The pilot system contained an image memory of 3M byte, which corresponds to a data storage of 0.3 sec duration at a rate of 2000 frames per second with 100*32 pixels per frame. In order to overcome this limitation, we have developed a special large-size digital image recording system. Using the new system, a preliminary observation of the vocal cord vibration for word initial consonants was conducted and the relationship between the pattern of the vocal cord vibration and the source spectrum was examined.

This paper presents the characteristics of the new digital image recording system using a fiberscope and the preliminary results of observation of the vocal cord vibration for the consonants.

New Digital Image Recording System

The new system was constructed as a special stand-alone digital video system. The number of picture elements can be selected from the following conditions and the frame rate of the image recording varies according to this choice.

200 * 64	500 frame/sec
200 * 32	1000 frame/sec
100 * 32	2000 frame/sec
100 * 16	4000 frame/sec

The system is equipped with a "finder" function. During the "stand-by" mode, the image can be monitored on the LCD display, which is refreshed at a reduced frame rate slower than the actual scan rate of the image sensor by sampling the video signals from the image sensor at intermittent frames. Stored images are reproduced as NTSC video signals. Reproduction of the image can be performed as with ordinary video tape recorders and is equipped with the following operation modes: PLAY, FAST-FORWARD, REWIND, SLOW-MOTION, STILL, REPEAT.

A personal computer can be connected to the system and image data from a selected part of the utterance can be sent to the computer for later data analysis.

Vocal Cord Vibration at Voice Onset in the Consonants

/t/ and /d/

Figures 1 and 2 show the vocal cord vibration at the utterance initial voice onset in /te:te/ and /de:de/, respectively. In the case of /t/, the first vibration takes place at around the frame A in Figure 1. In this cycle, the glottal closure is not complete and the posterior end of the glottis appears to remain open. In the next cycle of vibration, closure is nearly complete.

In the case of /d/, the left and right vocal cords reach approximated state before the start of vibration and stay in that position for a short period. Then, at around frame A in Figure 2, the glottis shows a slight opening as the first vibratory movement. This glottal opening is followed by a relatively long closed period compared to that in the following steady periods.

Figure 3 shows the inverse filtered source wave and the source spectrum for the first few cycles in /te:te/ and /de:de/. It can be seen in Figure 3, that, in the case of /t/, the spectrum for the 1st period of the source wave has a sharp decrease in the high frequency components which can be associated with the incomplete closure of the glottis. There is a large difference in the source spectrum between the first cycle and the following cycles. Compared to this, in the case of /d/. the spectrum of the first cycle is close to that of the following cycles. The difference between /t/ and /d/ in the source spectrum can be taken to reflecting the difference in the mode of the initiation of the vocal cord vibration observed in Figures 1 and 2.

/h/ and /s/

Figure 4 shows the start of vocal cord vibration in the word initial /h/ in /he: he:/.

The first glottal approximation takes place at around the frame marked A in the figure. As can be seen in the figure, at this moment the posterior part of the glottis remains open, and even in the next cycle, the glottal closure is still incomplete.

Compared to this, in the case of /s/ shown in Figure 5, complete glottal closure is generally observed in the second cycle, and a tight glottal closure is achieved rather quickly.

As indices of the glottal vibration, time series of the brightness curve at a selected horizontal scan line on the glottal image are displayed in Figure 6. A scan line close to the posterior end of the glottis is shown in this data display. Dips in the brightness curve towards the right correspond to the dark area of glottal opening. This figure shows that, in the case of /h/, the effect of vibration can be seen up to the eighth cycle after voice onset, whereas in the case of /s/ effect of vibration disappears after the second cycle.

This difference in the pattern of glottal vibration between /s/ and /h/ is clearly reflected in the spectrum of the voice source. Figure 6 also shows the spectrum of the inverse filtered speech waves for the several pitch periods after voice onset. In the case of /s/, the spectrum in the second pitch period already reaches a the same stationary pattern as that in the following vowel. In contrast to this, for /h/, the spectrum change continues over 5-6 pitch periods.

In summary, the present study of fiberscopic high-speed digital imaging of vocal cord vibration revealed that in the case of /h/ vibration starts with a wider separation of the vocal cords. The adductive movement of the vocal cords continues over several pitch periods after voicing, and there is a corresponding change in the source spectrum.

Reference

(1) K. Honda, S. Kiritani, H. Imagawa and H. Hirose: Laryngeal Function in Phonation and Respiration, Acollege-Hill Publication, 485 (1987).

(2) H. Imagawa, S. Kiritani and H. Hirose: Japanese J Medical Electronics and Biological Engineering, 25, 284-290, 1987.

(3) S. Kiritani: Vocal Cord Vibration and Voice Source Characteristics -Observations by the High-speed Digital Image Recording-, Proc. 12th ICPhS, 1, 206-209, 1991.

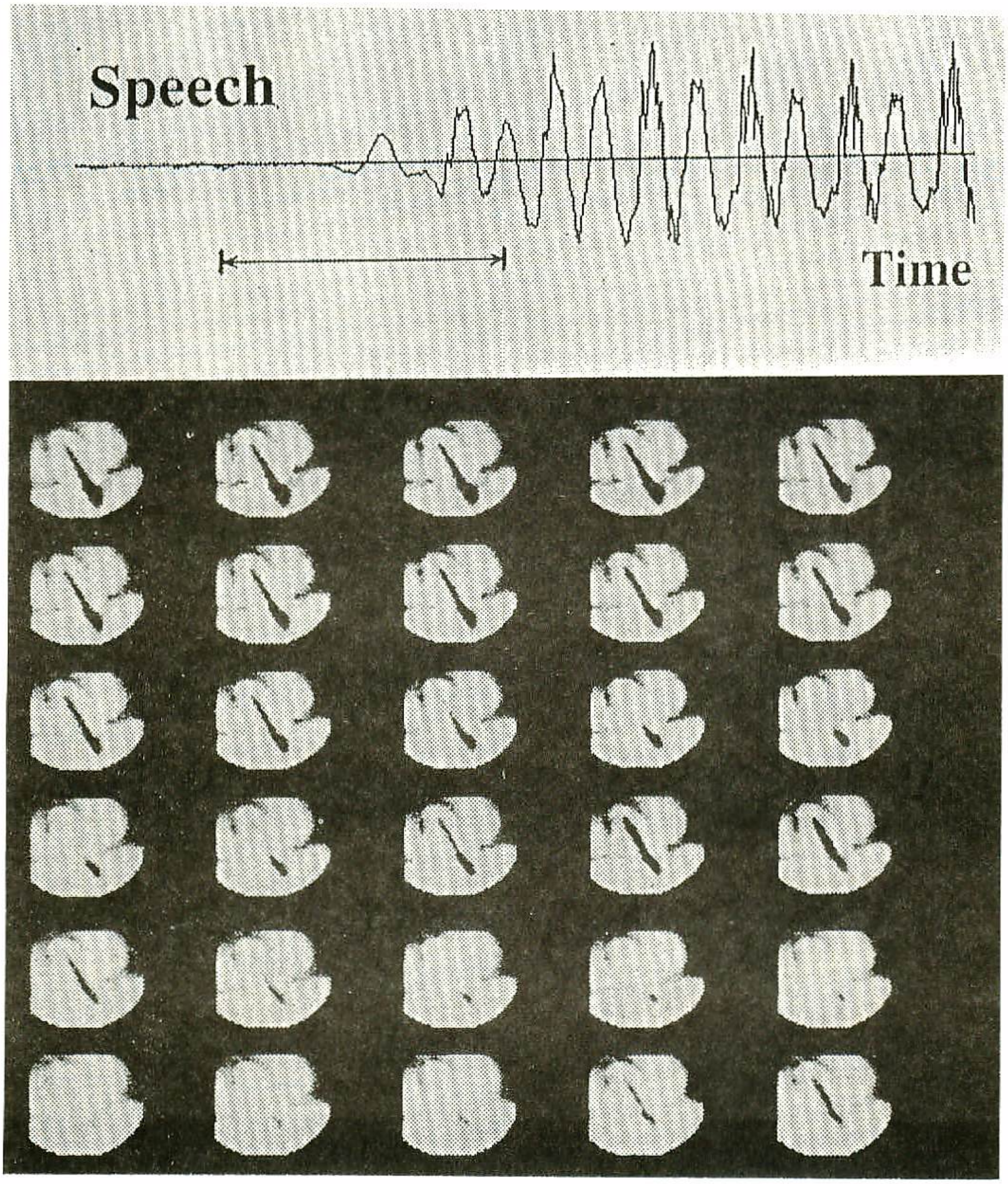

Figure 1. Vocal cord vibration at voice onset in /te:te/.

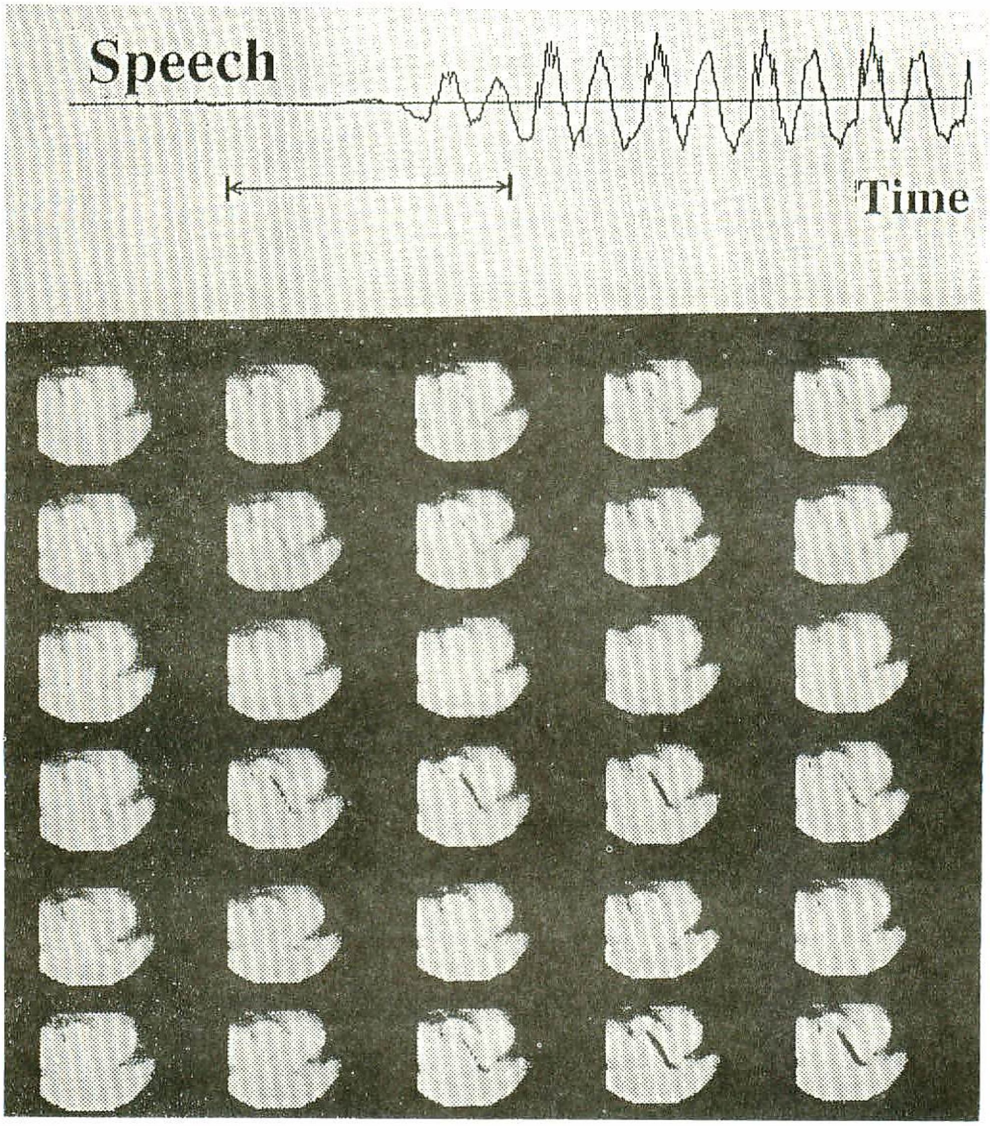

Figure 2. Vocal cord vibration at voice onset in /de:de/.

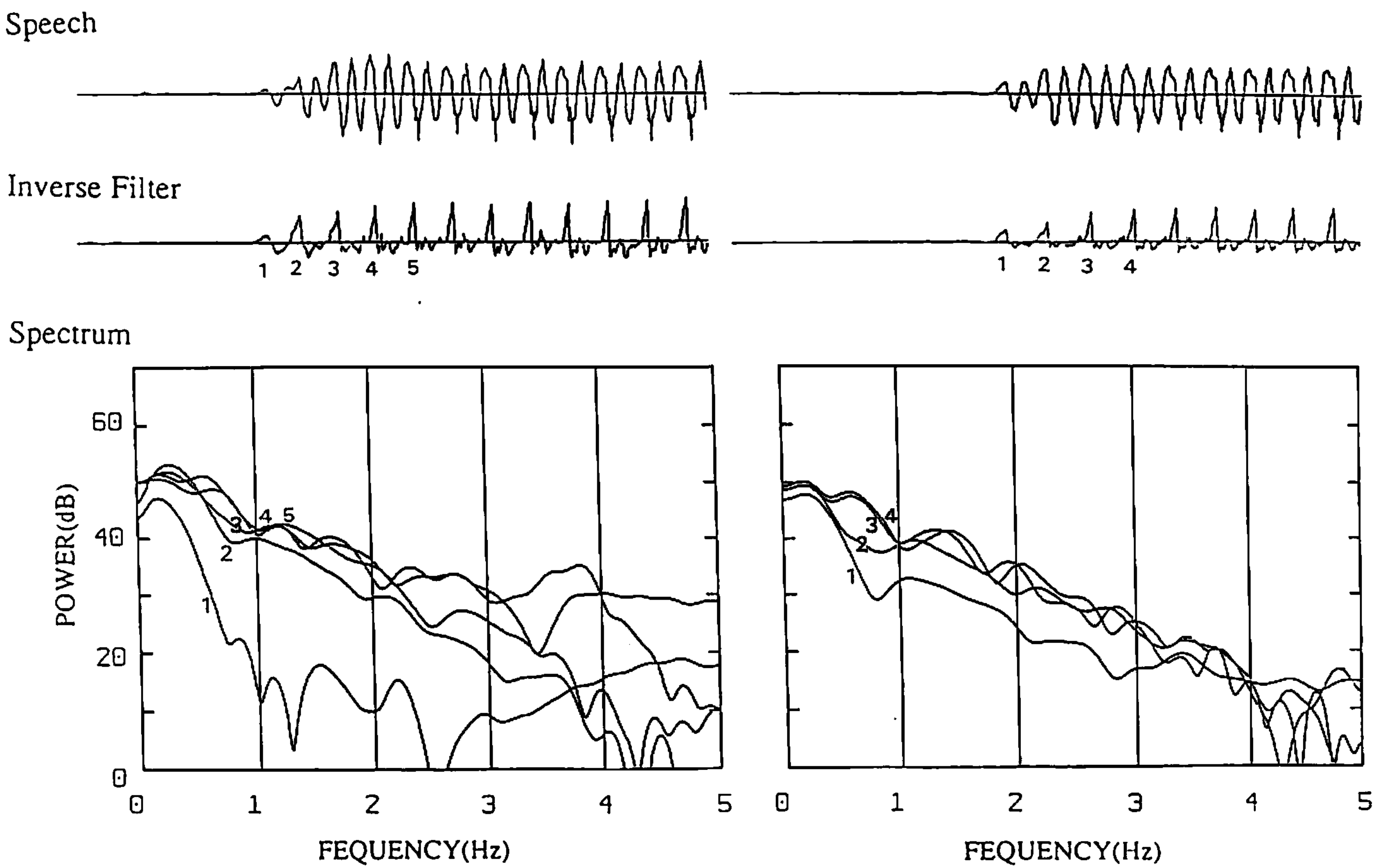

Figure 3. Inverse-filtered source wave and spectrum in /te:te/ and /de:de/.

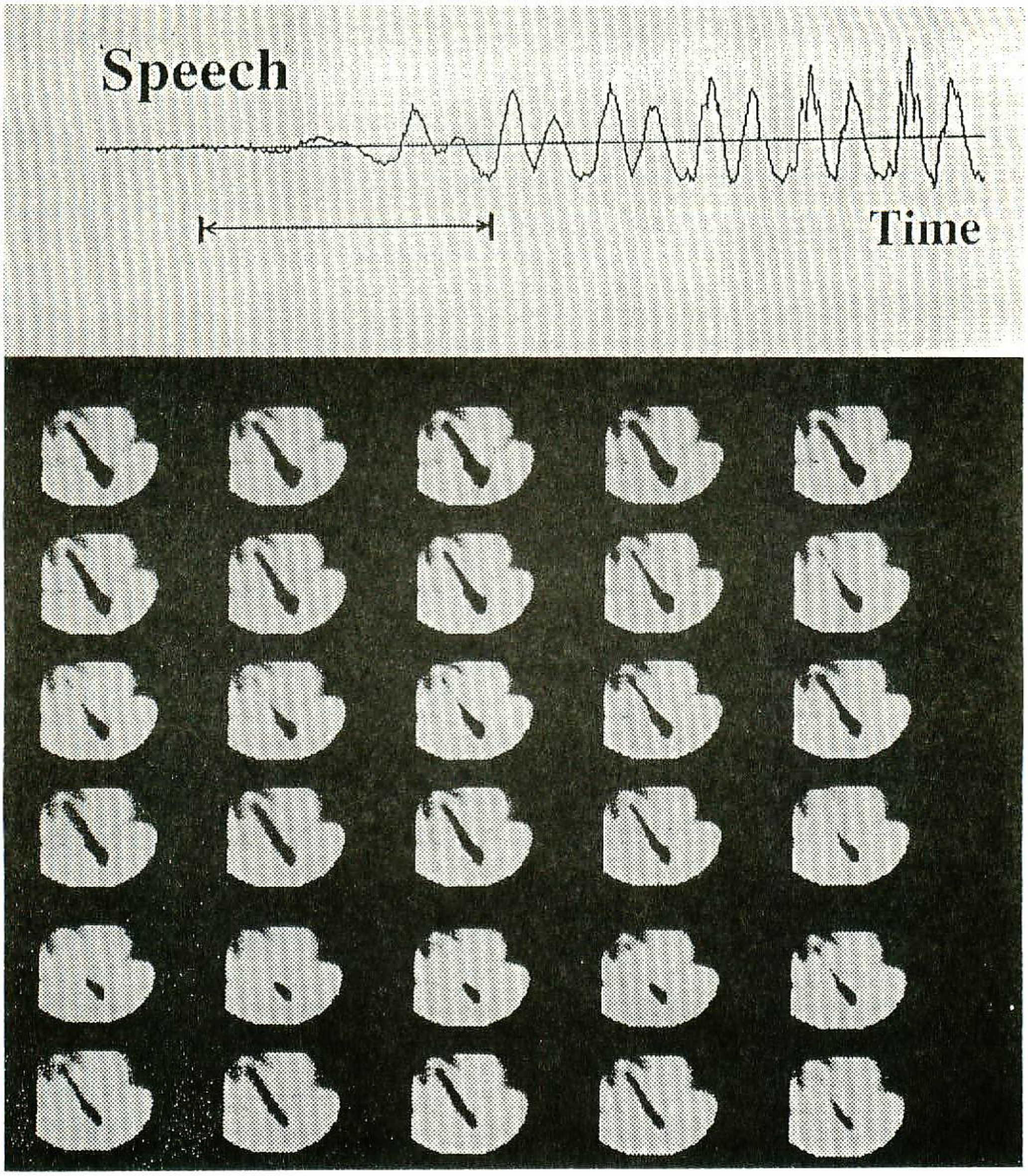

Figure 4. Vocal cord vibration at voice onset in /se:se/.

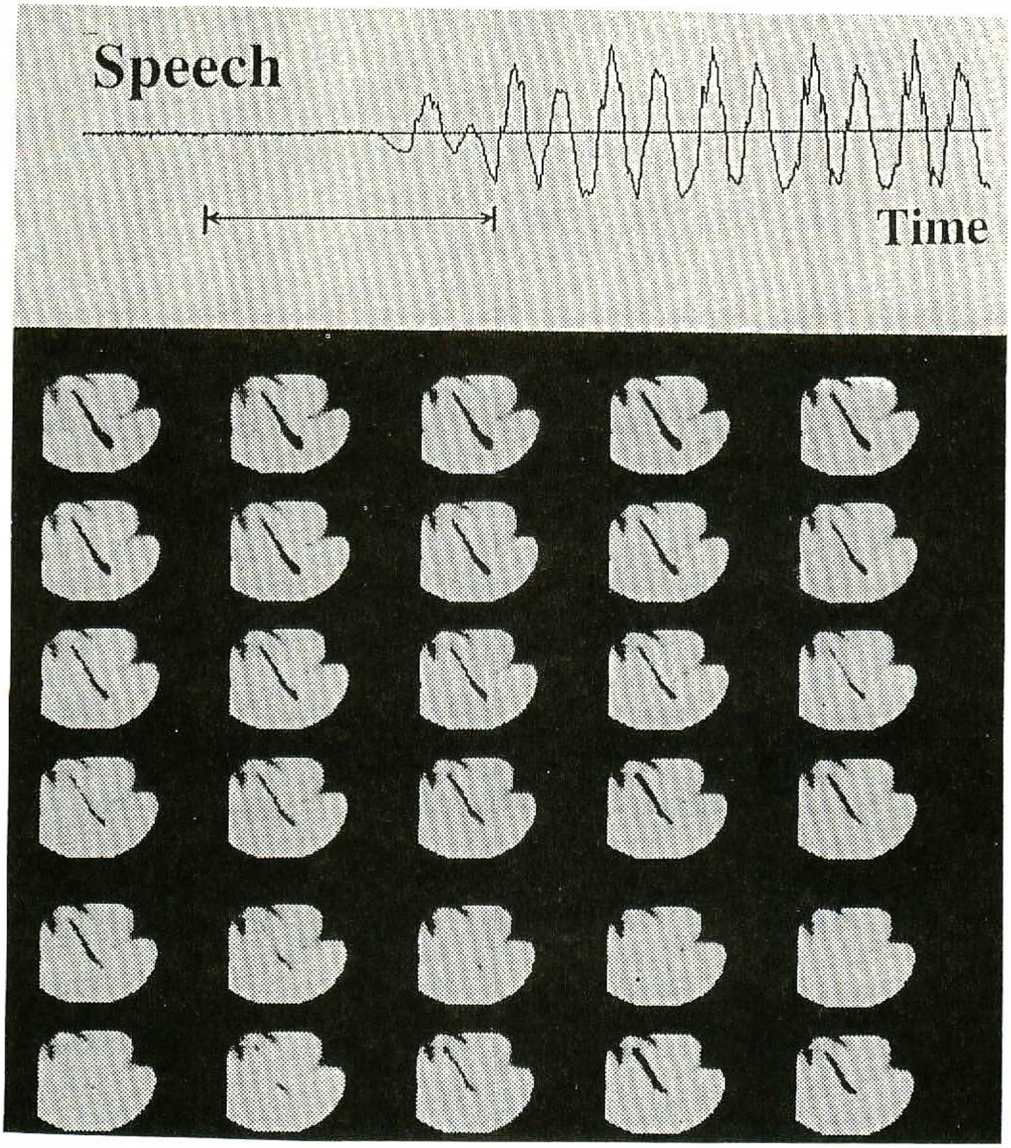

Figure 5. Vocal cord vibration at voice onset in /he:he.

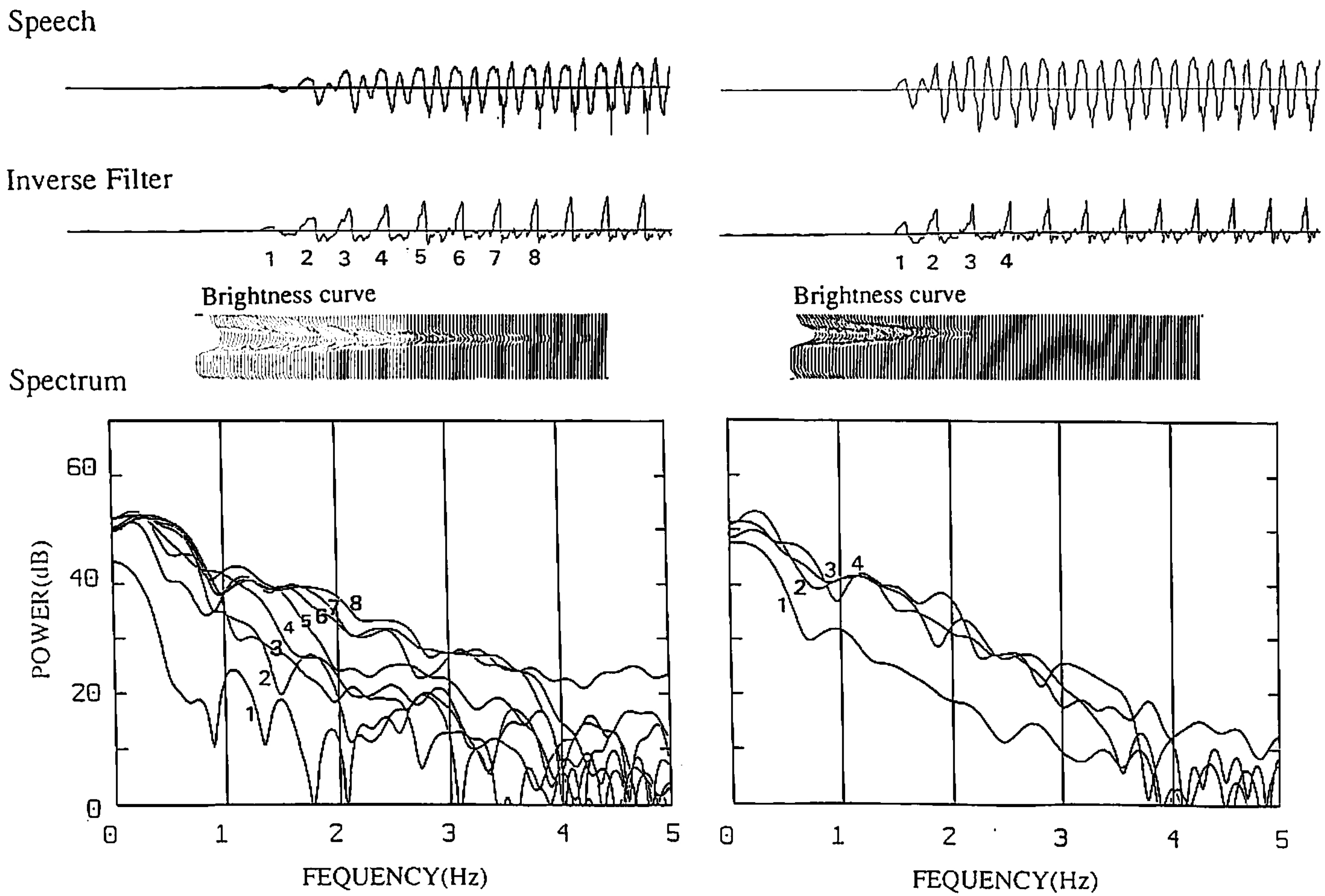

Figure 6. Inverse-filtered source wave and spectrum in /he:he/ and /se:se/.