VOCAL CORD VIBRATION AND VOICE SOURCE CHARACTERISTICS --OBSERVATIONS BY A HIGH-SPEED DIGITAL IMAGE RECORDING--

S. Kiritani, H. Hirose and H.Imagawa

Res. Inst. Logopedics and Phoniatrics,
University of Tokyo, Tokyo, Japan.

ABSTRACT

By using a high-speed digital image recording system, the relationship between vocal vibration and voice source characteristics has been investigated. Fiberscope was used to observe vocal cord vibrations during running speech. Pattern of vocal cord vibration at the onset and offset of the consonants were analysed. Solid endoscope system was used for observing sustained phonation of pathological voices. Many cases of rough voice show asymmetric and/or asynchronous movements between the right and left vocal cords, and between the anterior and posterior parts of the vocal cords. These movements appear to be related the periodical fluctuations in the vibratory pattern.

1. INTRODUCTION

For the study of the voice source characteristics, it is essential to record the vocal cord vibration simultaneously with the speech signal and to analyze the relationship between the pattern of the vocal cord vibration and the acoustic characteristics of speech signal. Observation of the vocal cord vibration has generally been performed by using a high-speed motion picture. However, that method requires special equipment and is not suited for flexible data collection.

In order to facilitate high-speed recording of vocal cord vibration, a special high-speed digital image recording system was developed by the present authors. The system consists of an image sensor and a digital image memory combined with the solid endoscope or the fiberscope. The system is small and compact and, thus, enables flexible data collection.

2. SOLID ENDOSCOPE SYSTEM

Fig.1 shows a block diagram of the system. The system consists of an oblique-angled solid endoscope, a camera body containing an image sensor and an image processor. The output video signal from the image sensor is fed into the image processor through a high-speed A/D converter. Stored images are then displayed on a CRT monitor as a slow motion picture, etc.

In order to obtain a brighter image, a new model of the solid endoscope was constructed. The diameter of the scope was larger than that of the

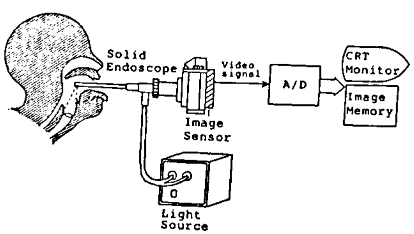

Fig.1 Blockdiagram of the solid endoscope system.

ordinary scope and contains two separate bundles of light guide. Light sources are the two 250W halogen lamps. Number of the picture elements in the image sensor is 100x100 and the sampling rate is 10MHz. In order to realize a high frame rate, a special scan method was devised in which only the selected scan lines were sampled. When 37 scan lines are sampled out of 100 scan lines, the frame rate is 2000 per second. The image memory is 1MByte and can store the image data for the period of about 200msecond.

3. FIBERSCOPE SYSTEM

In order to perform observation of the vocal cord vibrations during consonants in running speech, a high-speed image recording system using a fiberscope was also developed. A special fiberscope was also constructed the diameter of which was slightly larger than that of the ordinary scope. At the same time, a CCD image sensor was employed in this system, because the image by the fiberscope is darker than that by the solid endoscope. The sensitivity of the CCD image sensor is generally higher than that of the MOS image sensor which was used in the solid endoscope system. The light source is a 300W xenon lamp. A frame rate of 2000 per second was achieved with the picture element of 112x32.

Because the main objective of the fiberscope system is the observation of the running speech, it is necessary to record high-speed images during utterances as long as several seconds. Thus, a special stand-alone digital video system was constructed. The system consists of 64MByte image memory which can store the image data of 6.5 second duration under the frame rate of 2000 per second. The system generates standard NTSC video images. It can be operated as an ordinary video tape recorder, and is equipped with following operation modes; PLAY, FAST-FORWARD, REWIND, SLOW-MOTION, STILL, REPEAT. A personal computer can be connected to the system and the image data of the selected part of the utterance were sent to the computer for the later data analysis.

4. VOCAL CORD VIBRATION FOR CONSONANTS

By using a fiberscope system, vocal cord vibrations during the production of the intervocalic consonants were observed.

Figure 2(a) shows vocal cord vibrations during the release of word initial /p/ in the utterance /i: pepe desu/. Up to about 20th frame in the figure, the glottis remains open and there is no appreciable change in the glottal opening. At about 20th frame, narrowing of the glottis starts and the vocal cords start to vibrate. In the first cycle of vibration, glottal closure is incomplete. However, in the next cycle, closure of the glottis becomes complete. It can be seen in the figure that in the subsequent cycles, 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 following vowel. This phenomenon contributes, at least in part, to the higher pitch frequency in the post-consonantal period. It is also noted that the start of the clear excitation in the speech wave corresponds to the appearance of the closed phase in vocal cord vibration. Contrary to this, during the implosion of consonant, the excitation pattern of the speech wave decays even when vocal cord vibrations still maintains complete closure. It appears that the decay of the excitation pattern at this phase is due to the formation of the closure in the vocal tract.

Figure 2(b) shows vocal cord vibration for the production of /b/ (word medial /b/ in /i: bebe desu/). Vocal cord vibration during the consonant closure and during the preceding vowel are compared. 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. The closed phase is even longer for /b/.

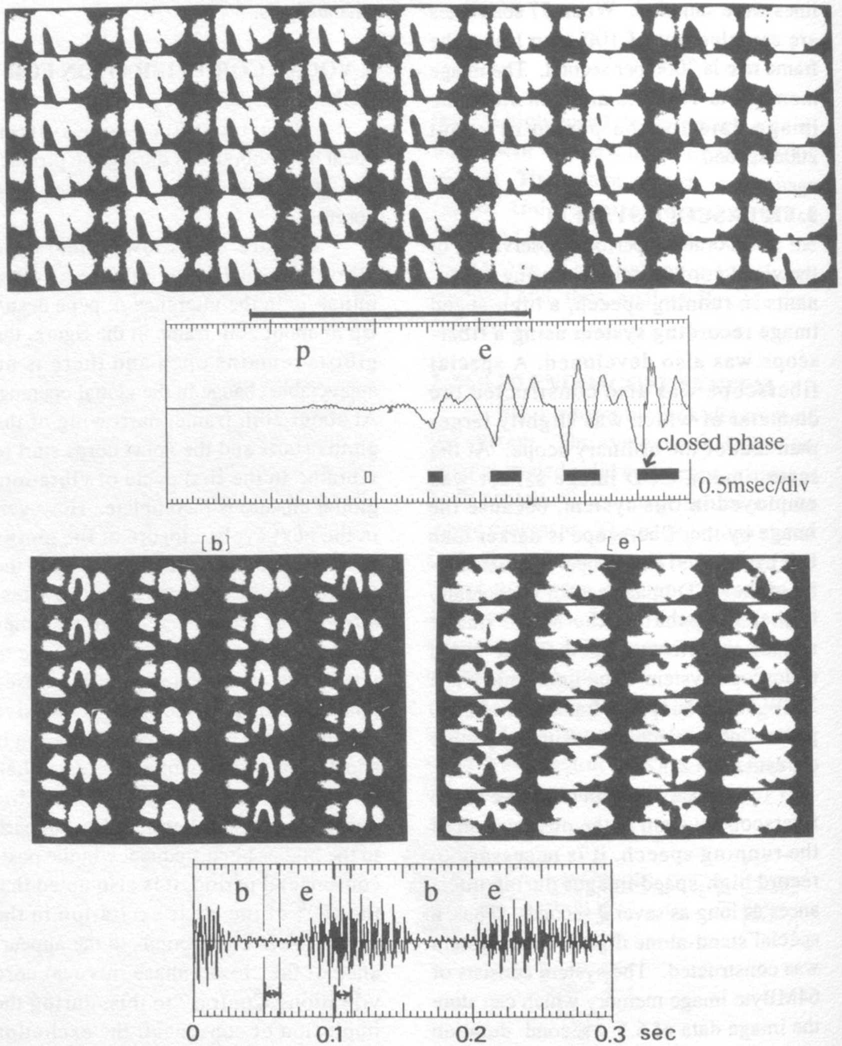

Fig.2 Vocal cord vibration in the production of the stop consonants.
(a) /i: pepe desu/ (b) /i: bebe desu/ Time: top to bottom, left to right.

5. VOCAL CORD VIBRATION IN ROUGH VOICES

The solid endoscope was applied to the observation of pathological vocal cord vibration in rough voices. Generally, rough voices have cycle to cycle fluctuations in speech waveform. However, many cases of rough voices show characteristic pattern of fluctuations (i.e., not purely random fluctuations). Fig. 3(a) shows a case of vocal fold cyst. The speech wave shows alternation of the period of strong excitation and the period of weak, noisy excitation. Glottal images in the figure reveal that there is a timing difference between the movements of the anterior and the posterior part of the vocal cord. The pattern of this timing difference varies from cycle to cycle. It appears that, in the period B, anterior part of the glottis starts to open as soon as the posterior part gets closed.

Fig. 3(b) shows a case of recurrent nerve palsy. In this case, the left and the right vocal cords show the difference in the vibratory frequency. During the preiod of high-amplitude waveform, the movements of the left and right vocal cords are in phase. In the following cycles, the left vocal cord gradually gets behind the right vocal cord and the movements of the vocal cords become out of phase. The vocal cords moves almost in pararell and the closed phase disappears. This state corresponds to the periods of the low-amplitude waveform. When the phase difference becomes greater than certain threshold, the vocal cords appear to pull each other and the movements again become in phase.

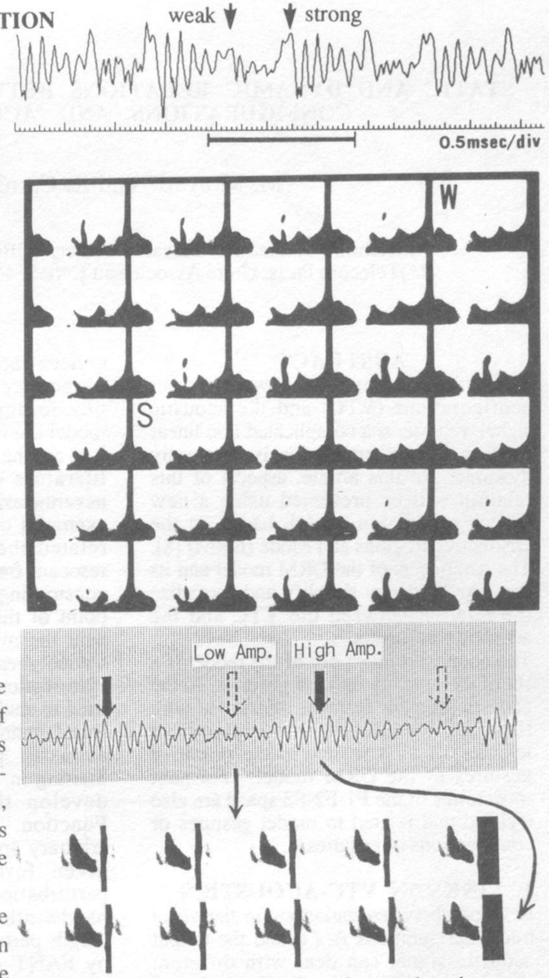

Fig. 3 Vocal cord vibrations in rough voices.
(a) Vocal cord cyst. (b) Recurrent nerve palsy.

These examples demonstrate that the present system is useful for the analysis of the production of the pathological voices.

1) S.Kiritani, K.Honda, H.Imagawa and H.Hirose: Proc. ICASSP, Tokyo, 1633 (1986).