

# VARIATION IN VOCAL FOLD VIBRATION ASSOCIATED WITH PROSODIC CONDITIONS

*Shigeru Kiritani, Yoshitake Iwamoto, Hiroshi Imagawa and Seiji Niimi*

Research Institute of Logopedics and Phoniatrics  
University of Tokyo, Tokyo, Japan

## ABSTRACT

Recently, a high-speed digital image recording of the vocal fold vibration at a rate of 4500 frames per second with the 256x256 picture elements has become possible. By using this technique, pattern of the vocal fold vibration under several different prosodic conditions, especially that of the utterance final position, were investigated. Imaging of the vocal fold vibration revealed that during the production of the vocal fry, the closure period from time to time to become very lengthy. After this lengthy closure period, the vibration tends to start with weak oscillation and during the following cycles, the oscillation builds up. It is suggested that this process explains the occurrence of multiple vibratory patterns associated with the vocal fry.

## 1. INTRODUCTION

The system of high-speed digital image recording developed by the present authors provides convenient means for the observation and analysis of the vocal fold vibration in various mode of phonation. The system is compact and is convenient for the simultaneous, synchronous recording of the speech signal, and thus is suited to the analysis of the relationship between the pattern of the vocal cord vibration and the acoustic characteristics of the speech signal. The system has been used at our institute for the studies of voice source characteristics in normal speech as well as in pathological voices[1-4].

Figure 1 shows a block diagram of the high-speed digital image recording system. The system can be used in combination with an oblique angled solid endoscope or with a flexible fiberscope. It consists of a camera body containing an image sensor and a digital image memory. The laryngeal image obtained through the endoscope or the fiberscope is focused on the image sensor. The image sensor is scanned at a high frame rate and the output video signal is fed into the image

memory through a high-speed A/D converter. Stored images are then reproduced consecutively as a slow-speed motion picture. In our previous system, we used a commercially available image sensor. In order to achieve a high frame rate, it was necessary to scan only a selected part of the sensor. The number of pixels was limited to 12632 to obtain a frame rate of 2,500 per seconds.

Recently, a new system with higher frame rate and resolution has become available [5]. The new system, combined with the use of an image intensifier, can also provide the fiberoptic image of a fairly good quality. By using this system, variation in the vocal fold vibration under several different prosodic conditions are now being studied. In this paper, examples of vocal cord vibration observed near the end of the natural utterances are presented.

**Figure 1:** Blockdiagram of the high-speed digital image recording system.

## 2. A HIGH-SPEED, HIGH-RESOLUTION IMAGE RECORDING SYSTEM

The new system employs an image sensor produced by the Photron Co. Ltd which was especially designed for the purpose of high-speed image recording. The sensor incorporates a technique of parallel read-out of image signals to obtain a high frame rate. The sensor contains 256256 picture elements and can be scanned at a rate of 4,500 frames per second. When the image area is restricted to 256128 picture elements, the frame rate is 9,000 per second.

An example of data obtained by this system (modal phonation by a male subject) is shown in Figure 2. Figure 2 (a) is a glottal image near the moment of the maximum glottal opening. In the present images, the edges of the vocal cords can be clearly identified by a

visual inspection. Figure 2 (b) shows movements of the edges of the left and right vocal cords measured on a selected horizontal scan line which is marked in the glottal image in (a). The figure 2 (c) shows successive frames at the moments of glottal opening and closing. These data shows that the present images by the present system can provide information nearly comparable to that obtained in the conventional high-speed motion picture studies.

## 3. IMAGE RECORDING USING A FIBERSCOPE

The image obtained through a flexible fiberscope is generally darker than that from the solid endoscope. When the fiberscope is directly connected to the present system, the brightness of the image is not sufficiently high for high-speed recording. However, the present system can be combined with an image intensifier and an image recording at a rate of 4,500 frames per second using a fiberscope is now realized. The image intensifier being used is the Channel Plate Type (C6276-01, HAMAMATSU Photonics Ltd.). The required light intensification factor is about 10. Because it is not so high, resulting image degradation due to the introduction of the image intensifier is minimal.

Figure 3 shows an example of the vocal cord vibration at the end of the natural utterance of the Japanese sentence, "Boku-wa ii" (I am OK). It was produced with a falling tone toward the end of the utterance. It can be seen in the figure that during the successive cycles, the amplitude of the vibration becomes smaller and when the vibration ceases, the vocal folds stay at a nearly adducted position. The vocal folds stay at that position for about

(a)

(b)

(c)

fluctuation is mostly due to the variation in the duration of the closure period. Compared to this, the variation in the duration of the open period is relatively small. From time to time, closure period becomes very long. It appears that in this phonation, vocal folds tend to be kept in the closed position.

In Figure 4 (b), the vocal fold vibration occasionally shows a bifurcate pattern. In these cases, the vibration in one cycle is weak in that the amplitude of the glottal opening is small and the glottal closure is incomplete. In the other cycle, the vibration is stronger and is followed by the complete glottal closure. However, the irregularity in the vocal fold vibration in this phonation is not limited to the bifurcate pattern (double vibration). In this phonation, the closure period from time to time

**Figure 3:** Vocal fold vibration at the cessation of voicing.

150 msec and then clear abductive movements start. Thus, in this phonation, the cessation of voicing is not directly caused by the abduction of the vocal folds. It appears that some other factors such as the tensening of the vocal folds and/or the change in the subglottal condition are responsible for the cessation of vocal fold vibration.

Another point to be noted in Figure 3 is that there is a difficulty in identifying the last several cycles of vibration on the EGG curve. This is presumably due to the incomplete, weak glottal closure in these vibrations which do not produce clear modulation of the EGG signal. The data indicates that in this type of phonation, EGG signal is not always a sufficient indicator of the vocal fold vibration.

#### **4. VOCAL FOLD VIBRATION IN THE VOCAL FRY**

Figure 4 (a) and (b) shows vocal fold vibrations associated with the vocal fries. They were recorded in the two separate phonations. It can be seen in the figure that at the initial part of the phonation (a), the open period is 14 msec and closed period is 6 msec approximately. Thus, the fundamental frequency is low (about 70 Hz) and an open quotient is small, as is generally known for the vocal fry. This voice shows a large fluctuation in the pitch period. However, the

(a)

(b)

**Figure 4:** Vocal fold vibration during the production of the vocal fry in the two different phonation.

becomes lengthy. After the relatively long closure period, the vocal fold vibration generally starts with a weak oscillation. It appears that in the following few cycles, the oscillation build up and when the oscillation becomes sufficiently large, the vocal fold vibration shows a complete closure and falls into the lengthy closure period again. It can be seen in the figure that this phonation contains various degrees of multiple vibration (double, triple, quadruple and sometimes more).

## 5. SUMMARY AND COMMENTS

The present paper presents the pattern of the vocal fold vibration near the end of the utterances recorded by using a high-speed digital image recording system. The present study shows that direct observation of the vocal folds sometimes reveals the existence of the vocal fold vibrations which are unclear on the tracing of the EGG signal. Vocal fold vibration during the production of the vocal fry tends to show lengthy closure period. After the relatively long closure period, vibration starts with weak oscillation and the vibration appears to build up during the following few cycles. This pattern of the vocal fold vibration conforms with that reported by Whitehead et. al. [6]. It is speculated that, in the production of the vocal fry, the vocal folds tend to be kept at a closed position. This is presumably due to the strong adduction of the vocal folds and/or the weak driving force in the glottal opening. Further physiological experiments are required to clarify the characteristics of this type of vocal fold vibration.

## 6. REFERENCES

1. K. Honda, S. Kiritani, H. Imagawa and H. Hirose: "High-speed Digital Recording of Vocal Fold Vibration Using a Solid-state Image Sensor," in *Laryngeal Function in Phonation and Respiration*, College-Hill Publication, 485-491, 1987.
2. H. Imagawa, S. Kiritani and H. Hirose: "High-speed Digital Image Recording System for Observing Vocal Fold Vibration Using an Image Sensor," *J. Medical Electronics and Biological Engineering*, 25, 284-290, 1987.
3. S. Kiritani, H. Imagawa and H. Hirose: "Vocal Cord Vibration and Voice Source Characteristics---Observations by the High-speed Digital Image Recording," *Proc. ICSLP-90*, Kobe, Japan, 61-64, 1990.
4. S. Kiritani, H. Imagawa and H. Hirose: "High-speed Digital Image Analysis of Vocal Cord Vibration in Diplophonia," *Speech Communication*, 13, 23-32, 1993.
5. S. Kiritani: "Recent Advances in High-speed Digital Image Recording of Vocal Fold Vibration," *Proc. ICPhs, Stockholm*, Vol. 4, 62-67, 1995.
6. R. L. Whitehead, D. E. Metz and B. H. Whitehead: "Vibratory Patterns of the Vocal Folds During Pulse Register." *J. Acoust. Soc. Amer.*, 75, 1293-1297, 1984.