ANALYSIS OF VOCAL FOLD VIBRATION DURING REGISTER CHANGE BY HIGH SPEED DIGITAL IMAGING SYSTEM

43.70.Jt

Masanobu Kumada*; Noriko Kobayashi**; Hajime Hirose**; Niro Tayama***;
Hiroshi Imagawa***; Ken-Ichi Sakakibara#; Takaharu Nito***; Mamiko Wada##;
Shin'ichi Kakurai***; Chieko Kumada***; Seiji Niimi###

*National Rehabilitation Center for the Disabled, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan, Tel: +81-42-995-3100 ex. 6676, Fax: +81-42-995-0355, E-mail: kumada@rehab.go.jp
**Kitazato University, Kanagawa, Japan
***University of Tokyo, Tokyo, Japan
#NTT Communication Science Laboratories, Kanagawa, Japan
##Tokyo Metropolitan Rehabilitation Hospital, Tokyo, Japan
###International University for Health and Welfare, Tochigi, Japan

ABSTRACT

Physiological study of prosody is indispensable in terms not only of the physiological interest but also of the evaluation and treatment for pathological cases of prosody. In free talk, the changes of vocal fold vibration are found frequently and these phenomena are very important prosodic events. Purpose and Method: To analyze the vocal fold vibration at the register change as the model of prosodic event, using our high speed digital imaging system at a rate of 4,500 images of 256.256 pixels per second. Tasks: Sustained phonations containing register changes. Results: Two major categories were found in the ways of vocal fold vibrations at the register change. In one category, changes were very smooth. In another, changes were not so smooth with some additional events, such as the anterior-posterior phase difference of the vibration, the abduction of the vocal folds, or the interruption of the phonation. Discussion: For the study of prosody, our high speed digital imaging system is a very powerful tool by which physiological information can be obtained.

INTRODUCTION

Physiological study of prosody is indispensable: the knowledge of how prosodic events are realized physiologically by articulators and phonators are very important in terms not only of the physiological interest but also of the evaluation and treatment for pathological cases of prosody such as stuttering, spasmodic dysphonia, Parkinson syndrome, ataxic speech caused by dysfunction of cerebellum, etc..

In free talk, the change of vocal fold vibration will be found frequently. That is, changes of the frequency of the vibration (intentionally controlled intonation, unintentional overturn of the pitch, etc.), or changes of the manner of the vibration (intentional and unintentional stop of the phonation, change of the voice quality, etc.). These phenomena are very important prosodic events.

Consistent vibration of the vocal folds in sustained phonation can be observed by stroboscopy. But when the vibration is inconsistent or changing its manner in short term such as in free talk shown above, stroboscopy can not be used, but a high speed digital imaging system is very powerful tool for the observation of the vocal fold vibration (Kiritani et al., 1996).

In this paper to begin physiological study of prosody in terms of vocal fold vibration, we will analyze qualitatively the vocal fold vibration at the register change as the model of prosodic events.

METHOD

<u>High Speed Digital Imaging System</u>

The system mainly consists of a camera head and a digital image memory. Either a solid endoscope or a flexible fiberscope is attached to the camera head for laryngeal image recording (Figures 1 and 2).

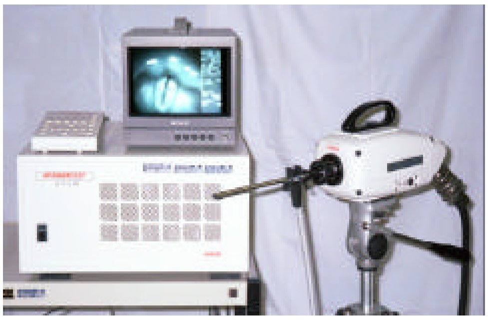

Figure 1 A picture of the camera head with an obliquely angled solid endoscope and a digital image memory with a video monitor. (Kiritani, 2000)

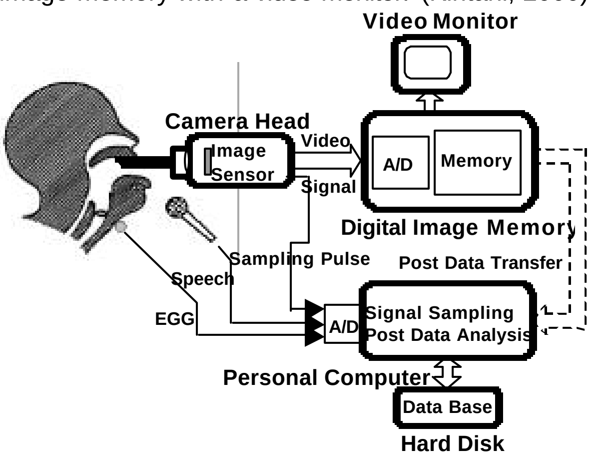

Figure 2. Block diagram of our high-speed digital image recording system using a solid endoscope. (Kiritani, 2000)

A solid endoscope combined with a 300W Xenon lamp light was used for this study for the observation of sustained phonation. (For observing vocal fold vibration during running speech, a flexible fiberscope is indispensable. Images obtained through a flexible fiberscope are generally darker than that obtained from a solid endoscope. Therefore, the system should be combined with an image intensifier. The required light intensification factor is not so much as 10, the resulting image degradation due to the use of the image intensifier is not significant.)

The camera head contains a solid-state image sensor specially designed for high-speed image recording. In order to achieve a high frame rate, the sensor with 256.256 picture elements incorporates a parallel read-out of image signals in its pixels (16 parallel channels).

Output video signals from the image sensor are fed into the digital image memory through A/D converters. The stored data also can be reproduced on a video monitor as slow-motion pictures. In this study with qualitative analysis, these slow-motion pictures (recorded on S-VHS videotapes and then transferred to PC as AVI files) were used.

The digital image memory can contain a maximum of 192 M Bytes of memory. Since each video signal is in 8-bit units (in 8-bit gray scale), this memory can store 3,000 frames of 256.256 images (=image data from 0.7 second phonation at a frame rate of 4,500 per second).

Data recording is usually done in a pre-trigger mode. When the system is set at “READY”, the video signal from the camera head is continuously sampled and stored in the image memory, which serves as a cyclic, endless-memory buffer. During this mode, the image on the video monitor is refreshed every 0.1 second. The image data just preceding the pressing of the "RECORD" button is memorized in the digital image memory.

In our system, not only the image data but also audio signals and electro-glottographic (EGG) signals are recorded simultaneously using a separate personal computer. To synchronize these data, sampling pulses generated by down-sampling the master clock pulse in the camera system are used. As the frame rate is 4,500 per second, four sampling pulses are generated in each image-recording time frame. Consequently, the sampling rate is 18 kHz. For qualitative data analysis, the image data stored in the image memory are transferred to the personal computer and stored in its disk memory, together with the digitized data from the speech and EGG signals.

Subjects

Four healthy adults who speak Tokyo dialects (two males (Subjects 1 and 2) and two females (Subjects 3 and 4)) were served as subjects for this study. The age average was 45.0 years (Sub.1: 31y.o., Sub.2: 68y.o., Sub.3: 27y.o., Sub.4: 54y.o.).

Tasks

Each subject was asked to try to phonate a Japanese vowel /e/ in head register and in chest register by turn (in one register after another in back-and forth manner) with no intermission of voicing between two registers; Each phonation was 2-3 second long and a register change was found every 0.5 second.

First, they tried to phonate in two registers with five degree scale (ex: D3 and A3 for Sub.1). All subjects except Sub.3 successfully could do it. Two series of images were recorded which contained a register change from chest to head. Other two series of images were also recorded which contained a register change from head to chest.

Second, they tried to phonate in two registers with octave scale (ex: C3 and C4 for Sub.1). All subjects successfully could do it. Again, two series of images were recorded which contained a register change from chest to head, and other two series of images which contained a register change from head to chest.

RESULTS

In qualitative analysis, two major categories were found in the ways of changing of vocal fold vibrations at the register change

In one category (Category A), changes were very smooth in terms of the vibration of the vocal fold vibration (Figure 3).

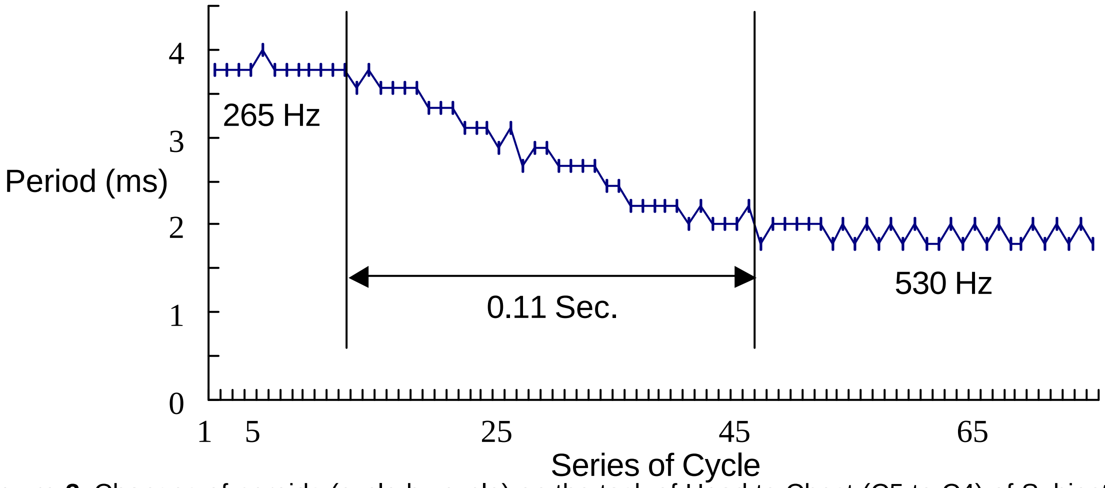

Figure 3. Changes of peroids (cycle by cycle) on the task of Head to Chest (C5 to C4) of Subject 4. It took 0.11 seconds to change register gradually (a typical example of Type A).

In another category (Category B), changes were not so smooth with some additional events at the register change, such as the anterior-posterior phase difference of the vibration (Type B-1, Figure 4), the abduction of the vocal folds (Type B-2), or the interruption of the phonation (in spite of the instruction) (Type B-3).

The summary is found in table 1.

DISCUSSION

For the manners of the changes of the vocal fold vibration at the register change, two major categories (A and B), and three subtypes for category B (Types B-1, B-2, and B-3) were found in qualitative analysis in this study.

Three subtypes for Category B were found in this study, however, the number of the subtypes for Category B is thought to increase if more subjects with more range of variety (ex. skill of phonation) are analyzed.

Consistent vibration of the vocal folds in sustained phonation can be observed by stroboscopy. But when the vibration is inconsistent or changing its manner in short term such as in free talk, stroboscopy can not be used, but high speed digital imaging system is very powerful tool for the observation of the vocal fold vibration, which was shown in this study of the register change.

Physiological study of prosody is indispensable: the knowledge of how prosodic events are realized physiologically by articulators and phonators are very important in terms not only of the physiological interest but also of the evaluation and treatment for pathological cases of prosody such as stuttering, spasmodic dysphonia, Parkinson syndrome, ataxic speech caused by dysfunction of cerebellum, etc..

Table 1. Types of the manner of the change of vocal fold vibration at the register change. Task: H to C: head register to chest; C to H: chest register to head; 5 Deg.: five degrees, Types: A: smoothly changed; B-1: the anterior-posterior phase difference of the vibration; B-2: the

abduction of the vocal folds; B-3: the interruption of the phonation.

Subject 1 (31y.o. Male)

Task	C to H 5 Deg. D3 > A3	C to H 5 Deg. D3 > A3	H to C 5 Deg. A3 > D3	H to C 5 Deg. A3 > D3	C to H Octave C3 > C4	C to H Octave C3 > C4	H to C Octave C4 > C3	H to C Octave C4 > C3
Freq.(Hz)	115>235	133>298	298>160	256>128	133>251	125>256	256>125	220>117
Type	B-2	B-2	A	A	A(B-2?)	A(B-2?)	A	A

Subject 2 (68y.o. Male)

Task	C to H 5 Deg. E3 > B3	C to H 5 Deg. E3 > B3	H to C 5 Deg. B3 > E3	H to C 5 Deg. B3 > E3	C to H Octave C3 > C4	C to H Octave C3 > C4	H to C Octave C4 > C3	H to C Octave C4 > C3
Freq.(Hz)	170>250	165>250	250>165	240>162	138>270	132>262	262>128	265>130
Type	B-2	B-2	A	A	B-2	A	B-1	B-1

Subject 3 (27y.o. Female)

Task	C to H Octave G4 > G5	C to H Octave G4 > G5	H to C Octave G5 > G4	H to C Octave G5 > G4
Freq.(Hz)	394>780	394>780	788>394	735>394
Type	B-3	B-3	B-3	A

Subject 4 (54y.o. Female)

Task	C to H 5 Deg. C4 > G4	C to H 5 Deg. C4 > G4	H to C 5 Deg. G4 > C4	H to C 5 Deg. G4 > C4	C to H Octave C4 > C5	C to H Octave C4 > C5	H to C Octave C5 > C4	H to C Octave C5 > C4
Freq.(Hz)	256>394	262>394	394>262	400>262	262>501	262>525	501>251	525>256
Type	A	A	A	A	A	A	A	

In free talk, the change of vocal fold vibration will be found frequently. That is, changes of the frequency of the vibration (intentionally controlled intonation, unintentional overturn of the pitch, etc.), or changes of the manner of the vibration (intentional and unintentional stop of the phonation, change of the voice quality, etc.). These phenomena are very important prosodic events. Therefore, for the study of prosody, our high speed digital imaging system is a very powerful tool by which physiological information can be obtained.

REFERENCES

1. S. Kiritani, H. Imagawa and H. Hirose: Vocal Cord Vibration in the Production of Consonants-Observation by Means of High-speed Digital Imaging Using a Fiberscope. Journal of the Acoustical Society of Japan (E), 17, 1-8,1996.
2. S. Kiritani,: Present System of High-Speed Digital Image Recording for Observing Vocal Fold Vibration. In "CD in Commemoration of the Retirement of Prof. Shigeru Kiritani", Secretariat for Commemoration of the retirement of Prof. Shigeru Kiritani, 2000.