Use of an Optical Position-Sensitive Device for Recording Lip and Jaw Movements in Speech

Y. Sonoda
Kumamoto, Japan

1. Introduction

For the study of the speech production process it is highly important to observe articulatory movements during speech. Most of the studies on articulatory behavior have been done by using x-ray techniques. Obvious disadvantages of x-ray observations are the exposure of the subject to harmful radiation and the time needed to extract data to obtain displays of motion in time. A computer-controlled x-ray technique has been developed to provide time-motion signals on articulators (Kiritani, Itoh and Fujimura, 1975). However, its use still must be under the constraints that the dose be minimized, and any scheme without the use of x-rays is useful both independently and as a supplement to an x-ray method.

This paper introduces a new optical recording system for speech research (Sonoda and Wanishi, 1982) and presents some early data on lip and jaw movements as examples.

2. Measuring Principle

The recording system consists of an optoelectric position-sensitive device (PSD) and light-emitting diodes (LEDs).The PSD provides two electrical analog outputs, specifying the individual X and Y positions of input light spots from LEDs, which are attached to several selected points of articulators. The system provides simultaneous recordings of the upper lip, lower lip, and jaw movements, continuously and in real time, with a high accuracy and a very fast response.

The PSD, developed and manufactured by Hamamatsu Corp. of Japan, is composed of a planar-type PN silicon photodiode with very uniform resistive layers formed on both the top and bottom surfaces as shown in Fig. 1. Pairs of electrodes (X_1-X_2 and Y_1-Y_2) are provided on each surface of the P and N layers, to sense X- and Y-axis signals, respectively. When a light spot from an LED is focused via a lens system on the active surface of the PSD, the photo current, which is proportional to the light intensity, is induced in the depletion layer at PN junction. This current is separated into two groups of output current, which flow through P and N layers to and from the opposite electrodes on the X and Y axis in the figure. Each of these currents is divided

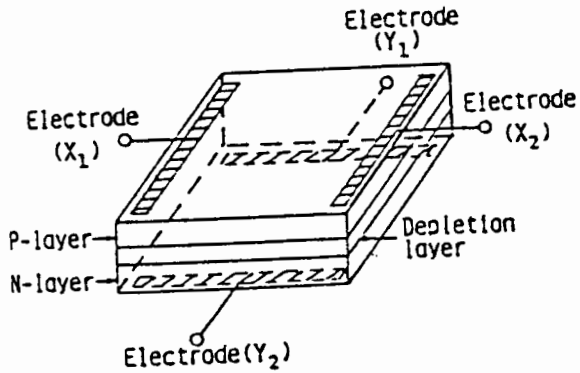

Fig. 1. Schematic construction of the PSD.

in accordance with the resistances of the corresponding layer between the light spot position and each electrode. Measuring each current and taking the ratio of their difference to their sum, the spot point on the PSD is converted into X and Y distance.

A circuit of the arithmetic operation utilized for a PSD is easily designed by the use of the analog IC arithmetic modules as shown in Fig. 2. The calibration patterns are slightly deformed around the edge. The maximum deviation in the useful area (5 × 9 cm) is estimated at 0.19 mm and the resolution on the position of an LED is estimated at 0.05 mm at a distance of 50 cm from the LEDs.

3. Data Collection

Preliminary experiments for recording lip and jaw movements have been conducted. Five of the eight LEDs were attached to these articulators within the sagittal plane as shown in Fig. 3. Two of the five LEDs were for monitoring the lip movements; upper lip (UL) and lower lip (LL). For the

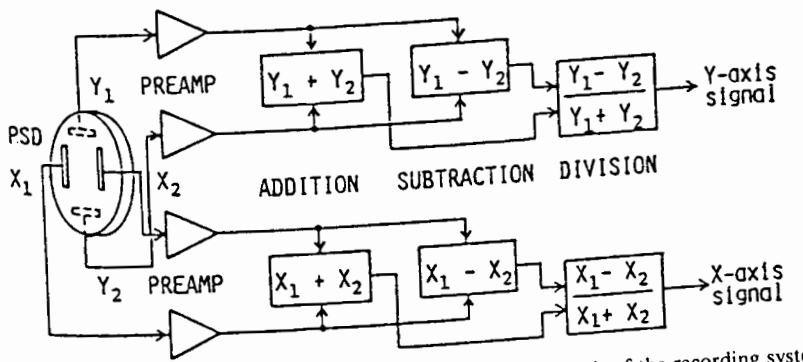

Fig. 2. Block diagram of the arithmetic positional detector unit of the recording system.

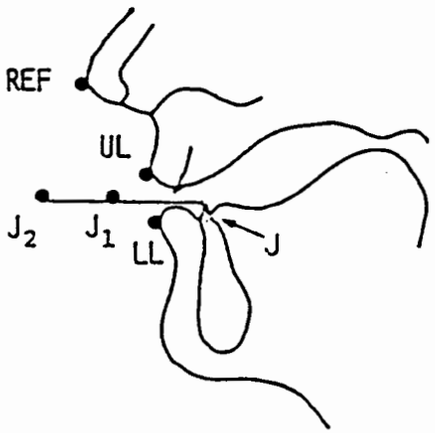

Fig. 3. Typical arrangement of LEDs used to observe articulatory movements.

observation of jaw movements, two LEDs (J_1 and J_2) were fixed on the rigid wire protruding from an artificial teethridge fixed at the canines. The movements of the jaw, the edge of the lower incisor J, were determined by these movements. The reference LED, REF, was used for monitoring the head displacement during experiments.

Fig. 4 shows an explanatory diagram for calculating the coordinates of the points J and C which correspond to the edge of the lower incisor and the point of the head of condyle, respectively. On these points hidden from view, the following calculations are made at each sample interval:

Jaw : $J(X_0,Y_0)$

$$X_0 = X_2+(X_1 - X_2)\frac{l_1 + l_2}{l_1}, \quad Y_0 = Y_2+(Y_1 - Y_2)\frac{l_1 + l_2}{l_1}$$

Condyle : $C(X_3,Y_3)$

$$X_3 = X_2 + \frac{1}{l_1}[(X_1 - X_2)l_0 + (Y_1 - Y_2)l_4]$$

$$Y_3 = Y_2 + \frac{1}{l_1}[(Y_1 - Y_2)l_0 + (X_1 - X_2)l_4], \quad \text{where } l_0=l_1+l_2+l_3$$

The parameters of l_3 and l_4 were obtained from the x-ray cephalograms for individual subjects.

In order to minimize gross movements during the experiments, the subject's head was fastened onto the headholder. However, small displacements of the head were observed and therefore subjects were instructed to take the jaw-closed position (rest position) keeping the lips relaxed after every tenth

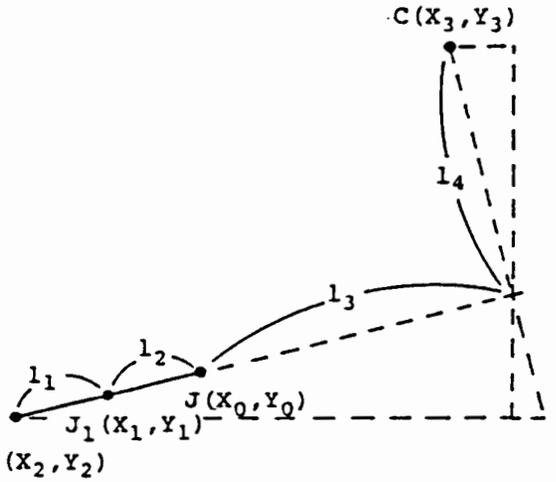

Fig. 4. Explanatory diagram for calculating the coordinates of the points J and C.

utterance of the test words. The data obtained under this condition were used for correction in terms of the translational and rotational perturbation in the head position during the course of an experiment. After the correction of the point REF to the origin, variations of the upper lip (UL) and jaw (J) were estimated. Standard deviation of the three subjects were very small; the individual values were 0.14, 0.52 and 0.43 (mm) at the point of upper lip, and 0.24, 0.44 and 0.41 (mm) at the jaw, correspondingly. These results show that the relative geometric relations among these points remained almost unchanged regardless of the small movements of the head. The lower lip is biomechanically linked with the jaw, the observed movement of the point LL is the sum of the individual lip and jaw components. The net movement of the lower lip was obtained by subtracting the movement of the point J from that of LL.

4. Experiments

All data were recorded once on an analog tape recorder and then stored in the computer at a frame rate of 200 Hz. Fig.5 shows typical time patterns (a) and two-dimensional trajectories (b) during an utterance of /epepe/ spoken by one subject. The curves in the figure are alternatively horizontal (H) and vertical (V) displacements of each point on the articulators. The curves denoted as L-J and CNDYL are the movement patterns of the lower lip itself and the head of the condyle respectively. The lowering of the upper lip for /p/ occurred synchronously with the closing (raising) of the lower lip. During the closure intervals for /p/, these were clipped and sustained movements on their position were observed in the figure. The jaw, however, gradually moved upward (closing) from the first vowel /e/ through the

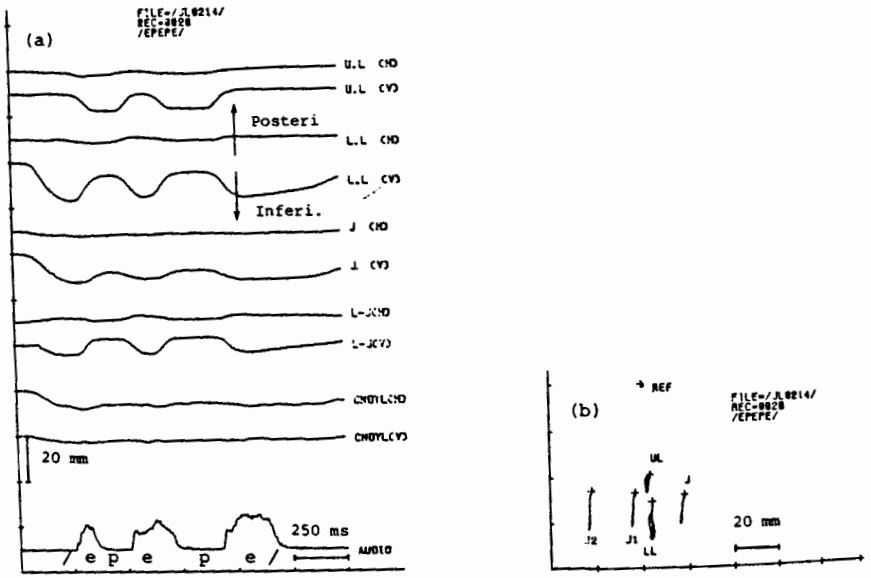

Fig. 5. Time patterns (a) and two-dimensional trajectories (b) of the articulatory movements in the utterance of /epepe/.

consonant /p/ and downward (opening) to the next /e/. During the first phase of jaw opening for /e/ from the closed (rest) position, the head of the condyle glided forward and somewhat downward. Since the consecutive movements of this point are very small compared with the corresponding jaw movements, the rotational movement of the mandible will apparently occur around the area on the head of the condyle.

The recording system presented in this paper may prove to be one of the most useful measuring techniques for the study of articulatory movements. By using this system, we are now conducting experiments to study the articulatory dynamics for several subjects.

References

Kiritani, S., Itoh, K., and Fujimura, O. (1975). Tongue-pellet tracking by a computer-controlled x-ray microbeam system. *J. Acoust. Soc. Am.* **57**, 1516-1520.

Sonoda, Y., and Wanishi, S. (1982). New optical method for recording lip and jaw movements. *J. Acoust. Soc. Am.* **72**, 700-704.