Ann. Bull. RILP, No. 10, 15-17 (1976)

PRINCIPAL COMPONENT ANALYSIS OF TONGUE PELLET MOVEMENT

S. Kiritani and H. Imagawa

For the parametric representation of articulatory movement, knowledge about the basic components of the tongue gestures in speech utterances is essential. In recent studies such as those by Ohta [1] and Shirai,[2] a method of multivariate statistical analysis of the articulatory movements was used to establish a set of effective articulatory parameters. In the former study, a principal component analysis was undertaken on the vocal tract area functions for Japanese vowels. These area functions were derived from speech signals using the Adaptive Inverse Filtering Method (Nakajima et al.).[3] In the latter study, the cineradiographic data on the tongue contours reported by Perkell was analyzed.

In the present study, principal component analysis was undertaken on the movement of tongue pellets tracked by an x-ray microbeam system. The articulatory data analyzed was obtained by tracking six pellets placed on the tongue, lower incisor, and lower lip as illustrated in Figure 1 (a). Figure 1 (b) shows the trajectories of the pellets for the pronunciation of a vowel sequence /ieaou/ in Japanese. The pellets were tracked at a rate of 120 frames per second. In the given pronunciation, pellets 3 and 4 on the tongue moved in a clockwise direction forming looped trajectories. Forward movement of pellet 2 was observed during the transition from /i/ to /e/.

The movements of five pellets (excluding the pellet on the lip) during the pronunciation of the vowel sequence /ieaouiaueo/ were subjected to the principal component analysis. A set of x- and y-coordinate values at each time frame constitutes one data point in 10 dimensional data spaces. A set of 360 time samples were analyzed.

The movements of the pellets corresponding to the individual principal components are shown in Figure 2. The first component mainly correlates with the front-back gesture of the tongue, the second component with the high-low gesture. The third component represents bulging of the tongue dorsum accompanied by the lowering of the tongue tip. Figure 3 shows pellet trajectories for /ieaou/ represented by using major two and three components. It is seen that the third component is necessary for describing the movement of pellet 2. Quantitative measures of the accuracy of describing observed pellet movement by selected major components are summarized in Table 1. The cumulative contribution rate of the first, second, and third components was 98%. The residual errors in the three-components approximation of pellet movement was 1 mm per pellet per frame. In Figure 4, the movement of the tongue for /ieaou/ is represented in a two-dimensional map using first and second components as the coordinate axes. It is noted that in this representation, movements from /i/ to /e/ and from /e/ to /a/ are approximately orthogonal. This orthogonal relation was not so clear in the earlier studies mentioned above.

Analysis of a larger set of data and corresponding acoustic signals are presently being conducted.

Table 1. Accuracy of principal component representation of pellet movement.

Principal component	Cumulative contribution rate	Residual error*
1	81%	3.5 mm
2	96%	1.5 mm
3	98%	1.0 mm

* per pellet/frame.

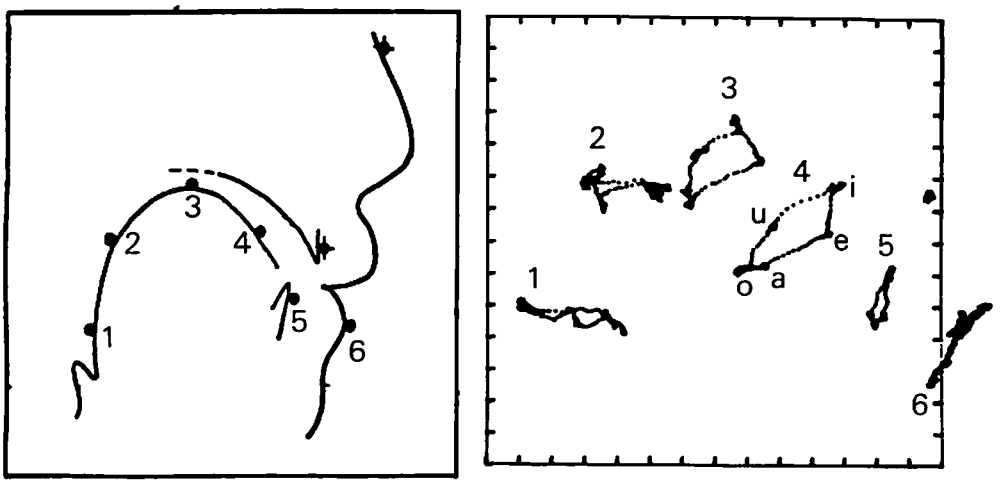

Fig. 1. (a) Pellet location for /ieaou/. (b) Pellet trajectories.

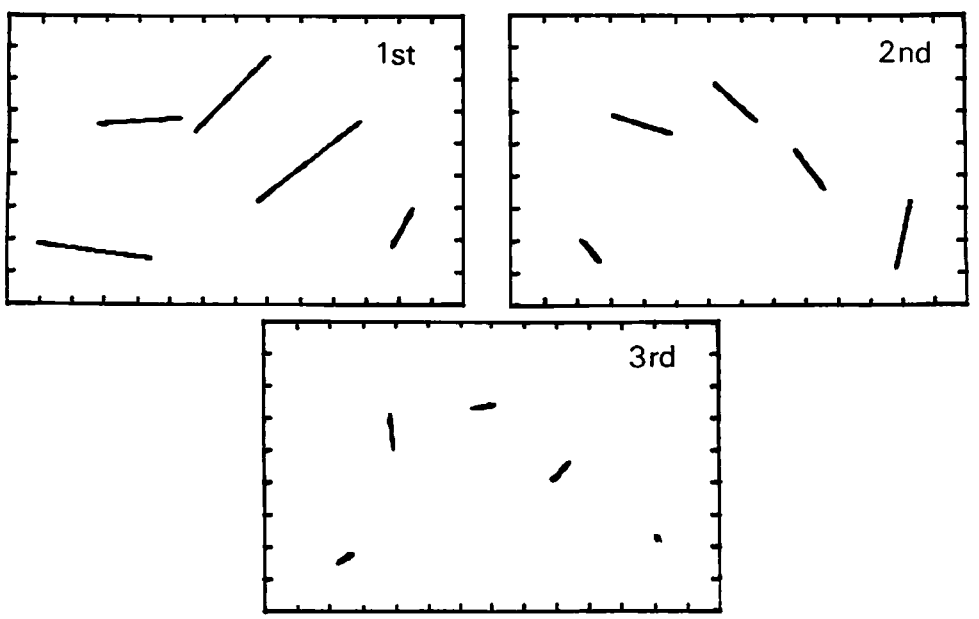

Fig. 2. Pellet movements corresponding to first, second, and third principal components.

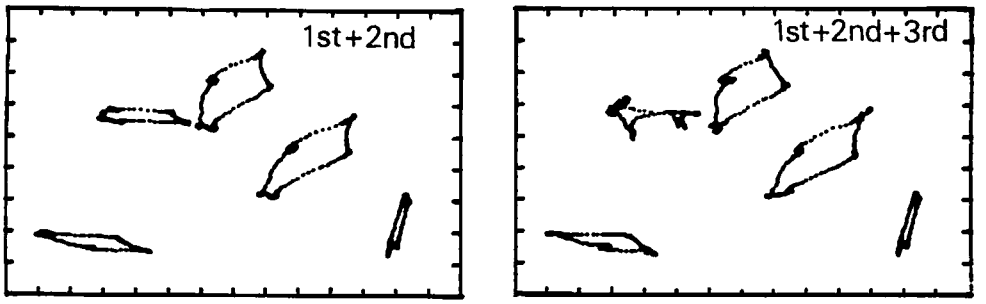

Fig. 3. Pellet trajectories for /ieaou/ represented by using major two and three components.

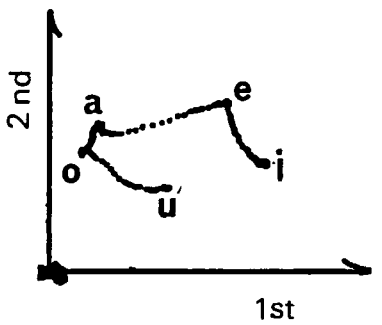

Fig. 4. Trajectory of tongue movement in the two dimensional principal component axis.

References

1. Ohta, K. (1974), "Studies of Estimated Vocal Tract Cross-sectional Area Functions of Japanese Vowels by Means of Multivariate Statistical Analysis," Transactions of the Committee on Speech Research, Acoustical Society of Japan, No. S74-27.
2. Shirai, K., H. Takagi, and T. Yoshioka (1974), "An Articulatory Model and the Estimation of Articulatory Parameters," Transactions of the Committee on Speech Research, Acoustical Sociey of Japan, No. S74-19
3. Nakajima, T., H. Ohmura, K. Tanaka and S. Ishizaki (1973), "Estimation of Vocal Tract Area Functions by Adaptive Inverse Filtering Methods," Bulletin of the Electrotechnical Laboratory, Vol. 37, No. 4, 462-482.