A SUPPLEMENTARY REPORT ON SPEECH PRODUCTION

Masayuki Sawashima, Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo, Tokyo, Japan.

In this paper, three different subtopics of speech production are discussed. They are: 1. Laryngeal control for voicing distinction. 2. Articulatory dynamics in normal and dysarthric cases. 3. Central mechanism of skilled movements.

1. Laryngeal Control for Voicing Distinctions

The basic features of laryngeal movements for the voiced vs. voiceless distinction in various languages have been examined by use of EMG and fiberoptic techniques. The glottal adduction-abduction dimension is directly observable with the fiberscope. The general picture is that the glottis is closed or nearly closed for voiced sounds while it is open for voiceless sounds, the extent of glottal opening varying with different phonemes and phonological environments. For Japanese voiceless stops the glottal opening for the same phoneme is greater in word-initial position than in word-medial position. For geminate stops, which occur only in word-medial position in Japanese, the duration of the glottal opening is consistently longer than for the corresponding non-geminate stops, whereas the degree of opening is often observed to be as small as in word-medial non-geminates. The findings for stops are also applicable to affricates. Voiceless fricatives show wide glottal aperture even in word-medial position. A large glottal aperture is associated with vowel devoicing in Japanese (Sawashima et al. 1976). In American English, a larger glottal opening associated with a greater degree of aspiration is observed for pre-stressed voiceless stops as compared to the corresponding post-stressed stops (Sawashima 1970). In French voiceless stops, a larger glottal opening is also observed for the pre-stressed position than for the post-stressed position (Benguerel et al. 1978). Observations on languages such as Korean (Kagaya 1974, Hirose et al. 1974), Hindi (Kagaya and Hirose, 1975) and Chinese (Iwata and Hirose, 1976), which have a phonemic distinction between aspirated and unaspirated stops, have revealed a large glottal opening for aspirated voiceless stops. The articulatory release takes place nearly at the point when the maximum glottal opening is reached. For unaspirated voiceless stops, on the other hand, the glottis is nearly closed to the phonatory position at the articulatory release, although a small amount of glottal separation is observable during oral closure. A similar contrast has been observed between /p/ and /b/ in Danish in word-initial position (Fischer-Jørgensen and Hirose 1974a). For the voiced aspirated stops of Hindi (Kagaya and Hirose 1975), the glottis is closed during most of the oral closure until it begins to open at the end of the oral closure, the maximum opening being reached after the release.

It has been observed that the glottal stop gesture, instead of glottal abduction, is used for English voiceless stops in certain environments (Fujimura and Sawashima 1971). The characteristic appearance of this gesture is an adduction of the false vocal cords covering the closed glottis. In whispered phonation, there is a constriction of the supraglottal laryngeal structures characterized by the adduction of the false vocal cords and a reduction in the anteroposterior dimension of the laryngeal cavity, although the glottis is open as in voiceless sounds in normal speech (Weitzman et al. 1976). The adduction of the false vocal cords appears to contribute to prevent the glottal vibration by the transglottal air flow and also to facilitate the generation of turbulent noise.

Electromyographic study of the larynx (Hirose and Gay 1972, Hirose and Ushijima 1978, Hirose et al. 1978a) has revealed, in various languages, a clear reciprocal pattern of activity between the posterior cricoarytenoid (PCA) and interarytenoid (INT) muscles for the voiced vs. voiceless distinction, a decrease in PCA activity with an increase in INT activity for voiced sounds, and the reverse for voiceless sounds. It has also been revealed that PCA is important for active vocal fold abduction for those speech sounds which are produced with an open glottis (Hirose 1976, Hirose and Ushijima 1978, Hirose et al. 1978a). A detailed observation of the laryngeal movements in correspondence with the EMG patterns for various types of Japanese voiceless sounds and sound sequences (Sawashima et al. 1978) has revealed that there is some subject-to-subject difference in the mode of laryngeal control using the PCA and INT muscles. In one subject, the time course and the extent of the glottal aperture are mainly represented by PCA activity with an associated decrease of INT activity. The time curve of the glottal width in this case can be interpreted as a kind of mechanically smoothed pattern of PCA activity. In the other subject,

however, the activity of the INT appears to actively contribute, in combination with the PCA, to the control of the glottal condition.

The data mentioned above present fairly clear physiological evidence for the laryngeal control of the glottal abduction and adduction. Another problem is whether or not we see physiological evidence for the stiff-slack dimension of the vocal folds, which was proposed by Halle and Stevens (1971) as another laryngeal feature contributing to the voiced-voiceless distinction in addition to abduction-adduction of the glottis. According to their proposal, stiffening of the vocal folds takes place for voiceless consonants and slackening for voiced consonants. When considering the physiological mechanism of control of vocal fold stiffness, we should refer to the "cover and body" structure of the vocal folds proposed by Hirano (1974). According to Hirano, the vocal folds consist of two different layers which are connected loosely with each other. The outer layer, which is called the "cover", is the mucosa covering the free edge of the vocal folds. The inner layer, which is called the "body", contains the vocalic muscle. The longitudinal pull of the vocal folds by the contraction of the cricothyroid (CT) muscle or some other external force results in an increase in the stiffness of both the cover and the body. Contraction of the vocalis (VOC) muscle also causes stiffening of the body, but it may shorten the vocal folds and result in slackening of the cover which would facilitate the vocal fold vibration (Fujimura 1977). Thus increase in the activity of CT can be physiological evidence for stiffening of the vocal folds, although other possible mechanisms are still to be explored. An extensive EMG study of the role of the larynx for the voicing distinction in Japanese consonants (Hirose and Ushijima 1978), has revealed that there is a temporary decrease in CT activity for both voiced and voiceless consonants. The degree of suppression is greater for word-initial voiced consonants than for the voiceless counterparts and least for word-medial consonants with no difference according to the voicing condition. There was also a temporary suppression of the VOC actitivy, the extent of the suppression being independent of the voicing condition but greater for the word-initial consonants than for the word-medial sounds. The results reveal that in Japanese consonants there is no physiological evidence for the stiff-slack dimension in the laryngeal control of the voicing distinction. The greater suppression of the muscle activities is considered to be related to the F_0 fall and the presence of the word boundary. In Hindi, Dixit (1975) reported a high CT activity for voiceless stops, but the results of Kagaya and Hirose (1975) for the same language failed to confirm that. A higher level of CT activity and a lower level of VOC activity for the voiceless stops than for the voiced stops is reported in a study of Swedish short and long consonants (Hirose 1977). Hirose et al. (1974) reported a sharp increase in VOC activity immediately before the articulatory release of a Korean forced stop. This particular VOC activity was interpreted as a physiological correlate of laryngealization as observed in the Danish stød (Hirose et al. 1974, Fischer-Jørgensen and Hirose 1974b, Fischer-Jørgensen 1977, Hirose et al. 1978a). Another interpretation proposed by Fujimura (1977, 1978) is that the vocalis muscle functions as a relatively fast-response voicing trigger mechanism for facilitating the vibration of the vocal folds which are otherwise under unfavorable conditions because of their tenseness.

In summary, physiological correlates of the tenseness feature appear to be manifested in some of the experimental results, but they are not as clear and consistent as those of the adduction-abduction dimension of the laryngeal features. It is reasonable to assume, however, that the laryngeal adjustments for the voicing distinction are not limited to simple adduction-abduction of the vocal folds. Further study is needed to explore the physiological correlates of some other features including the problem of tenseness of the vocal folds.

2. Articulatory Dynamics in Normal and Dysarthric Cases

In studying dynamic aspects of the articulatory movements, various basic characteristics such as the velocity of movement in different parts of the speech organs should be taken into account. Analysis of the articulatory movements in the repetitive production of a monosyllable is considered to present valuable information in this respect. According to Hudgins and Stetson (1937), the maximum rate of syllable repetition (mean value per second) is: 6.7 for the lip in /pu/, 8.2 for the tip of the tongue in /tat/, 7.1 for the back of the tongue in /ka/, and 6.7 for the velum in /tun/. Recordings of the actual movements and EMG of the relevant muscles provide data not only on the velocity but also on other aspects such as accuracy of the movement, regularity of the rhythmic pattern and muscle coordination. The development of the X-ray microbeam system (Kiritani et al. 1975) enabled us to make detailed analyses of the articulatory movements in normal subjects and also in patients with various neuromotor disorders (Hirose et al. 1978b, Hirose et al. 1978c).

The maximum velocity (mm/sec) in the syllable repetition for a normal subject is: 190-250 for the lip in /pa/, 220 for the tip of the tongue in /t/ of /pataka/, 200-220 for the back of the tongue in /ka/, and 105 for the velum in /teN/. It is noted that the velocity of the velar movement is definitely slower than the others. In the normal subject, the repetition of the movement is carried out quite regularly in terms of the amplitude, velocity, interval and also direction of the movement. In the normal subject it is also noted that attempts to make the syllable repetition at a slower rate do not result in a decrease in velocity but in an increase in the closure period of the given consonant as compared to a faster rate of repetition. EMG of the pertinent articulatory muscles shows a quite regular rhythmic pattern of activation-suppression corresponding to the movement with a clear reciprocal activity pattern between the antagonistic muscle pairs. The syllable repetition by patients with amyotrophic lateral sclerosis (ALS) is characterized by a slow rate of repetition and a decrease in both the velocity and amplitude of the movement, while the regularity of the movement is maintained. Patients with cerebellar ataxia are characterized by an irregular fluctuation of the interval, velocity and amplitude of the movement in syllable repetition. Electromyographic patterns also reveal irregularity in both the extent and timing of muscle activation. There is a plateau during the period of suppression indicating a disturbance of initiation of muscle activity in repetitive movements. Reciprocity between the antagonistic muscles is somehow preserved. The abnormal characteristic pattern of patients with Parkinsonism is: a small range of amplitude with a repetition rate comparable to normal. In addition, the amplitude gradually decreases throughout the repetition series until the movement finally stops. A gradual decrease in the velocity is also characteristic. Electromyographic records reveal a regular pattern of activation-suppression in each muscle, but the temporal reciprocity between the antagonistic muscles is not maintained and the two muscles are rather synchronously activated.

The dynamic characteristics presented here appear to well reflect the underlying motor pattern of voluntary movements in normal and various types of pathologic conditions. Thus the analysis of the syllable repetition is a promising approach for a differential diagnosis of various types of dysarthrias as well as for the study of dynamic aspects of speech production.

3. Central Mechanism of Skilled Movements

The central mechanism of dynamic adaptive motor control in speech production has been discussed by many researchers. One thing to be kept in mind here is the fact that the articulatory movements, although speech specific, are a kind of learned skilled voluntary movements. In this sense, it would be useful to refer to the central mechanism of other skilled movements which was suggested by Allen and Tsukahara (1974). They discussed the participation of the cerebellum in the planning and carrying out of a voluntary movement of the limbs in their extensive review on the cerebro-cerebellar communication systems. According to them, the most reasonable possibility for the lateral cerebellum is that it participates in the programming or long-range planning of the movement. The intermediate cerebellum works as a feedback system to the motor cortex in the execution of the movement. They state:

"Once the movement has been planned within the association cortex, with the help of the cerebellar hemisphere and basal ganglia, the motor cortex issues the command for movement. At this point the pars intermedia (of the cerebellum) makes an important contribution by updating the movement based on the sensory description of the limb position and velocity on which the intended movement is to be superimposed. This is a kind of short-range planning as opposed to the long-range planning of the association cortex and lateral cerebellum.... In learning a movement, we first execute the movement very slowly because it cannot be adequately preprogrammed. Instead, it is performed largely by cerebral intervention as well as by constant updating of the intermediate cerebellum. With practice,a greater amount of the movement can be preprogrammed and the movement can be executed more rapidly (without reference to peripheral sensory input." Thus, for learned movements the cerebellum provides an internal substitute for the external world. "This cerebellar operation we consider to take place in the lateral zone." Although they discuss only the control of limb actions, the main points may also be applied to the speech actions. In studying speech dynamics, we should refer to a more general physiological basis of development and organization of the skilled movements on one hand, and explore various speech specific problems on the other hand.

References

Allen, G.L. and N. Tsukahara (1974): "Cerebrocerebellar communication system", Physiol. Rev. 54, 957-1006.

Benguerel, A-P., H. Hirose, M. Sawashima and T. Ushijima (1978): "Laryngeal control in French stop production: a fiberscopic, acoustic and electromyographic study", Folia Ph. 30, 175-198.

Dixit, R.P. (1975): "Neuromuscular aspects of laryngeal control: with special reference to Hindi", Ph.D. Dissertation, the Univ. Texas at Austin.

Fischer-Jørgensen, E. (1977): Discussion in Dynamic Aspects of Speech Production, M. Sawashima and F.S. Cooper (eds.) 70, Univ. Tokyo Press.

Fischer-Jørgensen, E. and H. Hirose (1974a): "A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production", Haskins SR-39/40, 231-254.

Fischer-Jørgensen, E. and H. Hirose (1974b): "A note on laryngeal activity in the Danish "stød"", Haskins SR-39/40, 255-259.

Fujimura, O. (1977): "Control of the larynx in speech", Phonetica 34, 280-288.

Fujimura, O. (1978): "Physiological functions of the larynx in phonetic control", Paper at the IPS-77, Miami.

Fujimura, O. and M. Sawashima (1971): "Consonant sequences and laryngeal control", Ann. Bull. RILP 5, 1-6.

Halle, M. and K.N. Stevens (1971): "A note on laryngeal features", Q. Prog. Rep. Res. Lab. Electron., MIT 101, 198-213.

Hirano, M. (1974): "Morphological structure of the vocal cords as a vibrator and its vibrations", Folia Ph. 26, 89-94.

Hirose, H. (1976): "Posterior cricoarytenoid as a speech muscle", Ann. Otol. 85, 334-343.

Hirose, H. (1977): "Electromyography of the larynx and other speech organs", in Dynamic Aspects of Speech Production, M. Sawashima and F.S. Cooper (eds.), 49-65, Univ. Tokyo Press.

Hirose, H. and T. Gay (1972): "The activity of the intrinsic laryngeal muscles in voicing control: an electromyographic study", Phonetica 25, 140-164.

Hirose, H., C.Y. Lee and T. Ushijima (1974): "Laryngeal control in Korean stop production", JPh. 2, 145-152.

Hirose, H. and T. Ushijima (1978): "Laryngeal control for voicing distinction in Japanese consonant production", Phonetica 35, 1-10.

Hirose, H., H. Yoshioka and S. Niimi (1978a): "A cross language study of laryngeal adjustment in consonant production", Ann. Bull. RILP 12, 61-71.

Hirose, H., S. Kiritani, T. Ushijima and M. Sawashima (1978b): "Analysis of abnormal articulatory dynamics in two dysarthric patients", JSHD 43, 96-105.

Hirose, H., S. Kiritani, T. Ushijima, H. Yoshioka and M. Sawashima (1978c): "Patterns of dysarthric movements in patients with Parkinsonism", Ann. Bull. RILP 12, 73-85.

Hudgins, C.V. and R.H. Stetson (1937) cited from Suprasegmentals, I. Lehiste, MIT Press, 1970.

Iwata, R. and H. Hirose (1976): "Fiberoptic acoustic studies of Mandarin stops and affricates", Ann. Bull. RILP 10, 47-60.

Kagaya, R. (1974): "A fiberoptic and acoustic study of the Korean stops, affricates and fricatives", JPh. 2, 161-180.

Kagaya, R. and H. Hirose (1975): "Fiberoptic, electromyographic and acoustic analysis of Hindi stop consonants", Ann. Bull. RILP 9, 27-46.

Kiritani, S., K. Itoh and O. Fujimura (1975): "Tongue-pellet tracking by a computer-controlled X-ray microbeam system", JASA 57, 1516-1520.

Sawashima, M. (1970): "Glottal adjustments for English obstruents", Haskins SR-21/22, 180-200.

Sawashima, M., H. Hirose and S. Niimi (1976): "Glottal conditions in articulation of Japanese voiceless consonants", XVI Int. Congr. Logopedics and Phoniatrics, Interlaken 1974, 409-414, Basel: Karger.

Sawashima, M., H. Hirose and H. Yoshioka (1978): "Abductor (PCA) and adductor (INT) muscles of the larynx in voiceless sound production", Ann. Bull. RILP 12, 53-60.

Weitzman, R.S., M. Sawashima, H. Hirose and T. Ushijima (1976): "Devoiced and whispered vowels in Japanese", Ann. Bull. RILP 10, 61-80.