ON THE USE OF OROSENSORY FEEDBACK: AN INTERPRETATION OF COMPENSATORY ARTICULATION EXPERIMENTS

Joseph S. Perkell, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass. 02139, U.S.A.

A number of experiments have been performed on "compensatory articulation" with the aim of understanding more about speech motor programming. Several of these experiments have used bite blocks to constrain the mandible in abnormally open (or closed) positions while the subjects produced steady state vowels (/i/, /a/, and /u/) (cf. Lindblom, 1971; Lindblom, et al., 1977; Lindblom et al., in press; Gay and Turvey, these proceedings). The resulting formant patterns were measured at the first glottal pulse to avoid any possible effects of auditory feedback (which was not masked out). It was found that vowels produced with significantly abnormal jaw openings (i.e. 22-25 mm open for /i/) were essentially the same in quality as those produced normally by the same subjects. However when bite blocks were used in conjunction with oral topical anesthesia (Lindblom, et al., 1977) or with a combination of oral topical anesthesia and anesthesia of the temporomandibular joint (Gay and Turvey, these proceedings), subjects needed several attempts to produce appropriate vocal tract configurations and sound outputs. In the latter experiment, the application of oral topical anesthesia alone was not enough to impair subjects' ability to produce vowels appropriately.

Lindblom and his co-workers interpret their findings as support for the following view of the role of orosensory feedback. Tactile information from the labial and oral mucosa can be utilized in the motor programming of speech. Vowel "targets" may be encoded as [oro]sensory goals which reflect a neurophysiological encoding of area functions. These goals serve as a basis for the elaboration of motor commands by structures which "can generate appropriately revised motor commands on the basis of the feedback positional information available before onset of phonation" (Lindblom, et al., in press).

These results and their interpretations must be viewed with caution for a number of reasons. For example, a generous application of topical anesthesia to the oral and pharyngeal cavities can have a distracting effect on the subject (Lindblom, personal communication). Perhaps more importantly, a steady-state paradigm which allows the subject time to "organize" his response before presenting it may reflect functions which are not part of normal dynamic speech motor processes (cf. Leanderson and Persson, 1972; Abbs and Eilenberg, 1976). Nevertheless, the results are provocative enough to warrant further examination, particularly in light of a recent experiment on arm movements and another experiment on compensatory articulation.

Polit and Bizzi (1978) have performed an experiment in which 3 adult monkeys were trained to point to a target light with the forearm and hold the position for about 1 second in order to obtain a reward. The monkey could not see its forearm which was fixed to an apparatus that permitted only flexion and extension about the elbow in the horizontal plane. Performance was tested before and after a dorsal root section which eliminated somatosensory feedback from both upper limbs. In both intact and deafferented animals, the arm was unexpectedly displaced within the reaction time of the monkey, and in both cases the displacement of the initial arm position did not affect the attainment of the intended final steady-state position. These results suggested to the authors that a central program specified an equilibrium point corresponding to the interaction of agonist and antagonist muscles. A change in the equilibrium point leads to movement and attainment of a new posture.[1] However, it was also found that when the spatial relationship between the animal's arm and body was changed, the pointing response of the deafferented monkeys was inaccurate, and remained so even when visual feedback was allowed. In contrast, the intact monkeys were able to compensate within a few tries to the new position without visual feedback. This finding suggested that one major function of afferent feedback is in the adaptive modification of learned motor programs (Polit and Bizzi, 1978).

Following these authors' interpretation of their results, we might consider that in establishing the central program for the performance of the motor task (i.e., learning the task), the monkeys were incorporating a subconscious "knowledge" of the relationships between the target points with respect to the

(1) The existence of additional processes related to the dynamic aspects of movements is acknowledged, but not treated by Polit and Bizzi (1978).

apparatus and the muscular settings which would result in correct pointing. In doing so, the monkeys were calibrating the biomechanical properties of the system with respect to a particular frame of reference (i.e. orientation in space in relation to the body) with the use of somatosensory feedback from the system. When the frame of reference changed, only the monkeys with intact somatosensory pathways were able to "recalibrate" the central program to the new frame of reference.

This line of reasoning and the interpretations of Lindblom and his colleagues lead us to a possible, slightly more specific explanation of the compensatory articulation results. In the case of steady-state vowel productions, the frame of reference is defined as the configuration of the dorsal walls of the vocal tract and the position of the mandible. The target (or goal) consists of a vocal-tract area function as sensed by a complex pattern of sensory feedback from the vocal tract. Normally, to produce a steady-state configuration, the control mechanism has a choice of: 1) using a pattern of peripheral feedback to compare with one that has been learned in association with a particular area function and vowel quality, or 2) using a set of equilibrium levels of muscle excitations. These muscle excitation levels can be stored or computed on the basis of an overlearned knowledge of the vocal-tract geometry and biomechanical properties.

Now let us consider the three possible combinations of the use of anesthesia and/or bite blocks. With only (complete) anesthesia, the controller uses option 2. In other words, with a frame of reference which is assumed to be normal, the controller is still capable of specifying equilibrium muscle excitations which it "knows" will produce the correct area function. On the other hand with only the bite block, the controller uses option 1. The appropriate area function is produced by comparing peripheral feedback with the "known" pattern. With anesthesia and the bite block, neither option is available. The frame of reference has been changed. The absence of feedback about the new frame of reference precludes an *a priori* recomputation of appropriate equilibrium muscle excitations, and the absence of tactile feedback precludes a direct comparison with the known pattern. This last statement is reinforced by the results of Gay and Turvey in which only combined anesthetization of the oral mucosa and the temporomandibular joint (along with the bite block) rendered the subject incapable of producing the vowel correctly on the first try. The loss of joint sensation would eliminate the feedback about the frame of reference, needed for a recalibration of the central program, and the loss of sensation from the oral mucosa would preclude using such feedback directly in an error-minimizing feedback loop.[2]

The hypothetical use of afferent information to keep the controller informed about the frame of reference would be equivalent to the function proposed by Polit and Bizzi (1978) in the adaptive modification of learned motor programs. Presumably, the predictive simulation mechanism proposed by Lindblom, et al., (in press) also needs to use feedback in a similar way. It has been suggested by numerous investigators that learning a motor activity consists in part of substituting central programming for the use of peripheral feedback. This use of central patterning presumably incorporates an ability to adjust the parameters of the central program to account for changes in the frame of reference. In the case of speech, such changes correspond to speaking with a pipe clenched between the teeth, with the head tilted to one side, or resting ones' chin in his or her hand.

Gay and Turvey also found that /i/ and /u/ productions are affected by the combination of joint and topical anesthesia with a bite block while /a/ is not. This difference might be explained by their report that the topical anesthesia was applied to the oral cavity, where the acoustically most critical points of maximal constriction for /i/ and /u/ are located. (A given change in the dorsal-ventral location of the tongue surface will have a proportionally larger effect on the vocal-tract cross-sectional area at the point of maximum constriction than at other locations where the cross-sectional area is greater.) If the anesthesia did not exert a strong effect at the point of maximal constriction for /a/, we might expect it to be produced normally, with the use of feedback which is less consciously obvious but still available from that region. The importance of the pattern of contact at

(2) The fact that only topical anesthesia in combination with bite block was sufficient to impair vowel production in the subject of Lindblom, et al. (1977) might be due to individual differences or differences in the extent and depth of topical anesthesia.

the point of maximal constriction for the vowel is further suggested in lateral cineradiographic tracings (Netsell, et al., 1978) of normal and compensatory productions of the vowel /i/ for 3 subjects. For each subject the normal and compensatory dorsal tongue contours show considerable overlap and the overlap is most pronounced at or near the point of maximal constriction.

This brief analysis greatly oversimplifies the issues in a number of respects. It relies on a small amount of data. It overlooks the significant differences between deafferentation and the application of anesthesia as well as the unnatural nature of both experimental paradigms. And as we have mentioned, it deals with a steady-state task which may be quite different from anything actually found in speech. For these reasons, we must be very tentative in extending our interpretations to cover normal articulatory movements. However, on the basis of considering a number of additional aspects or speech production and the control of movement (see Perkell, 1979a), it is possible to offer the following speculations on the use of orosensory feedback.

1) Orosensory feedback may play a role in determining the nature of some distinctive features. It is possible that certain well-defined patterns of orosensory feedback (such as contact of the tongue with maxillary structures) facilitate the production of sounds which have distinctive acoustic and auditory-perceptual correlates (see also, Stevens and Perkell, 1977, Perkell, 1979b). Such patterns of orosensory feedback could be the speech production correlates of distinctive features. Specifications of utterances in the form of feature-related complexes of orosensory goals might serve as a basis for the production of articulatory movements. Thus, orosensory feedback on a long-term basis might be necessary for the establishment and maintenance of a subconscious "knowledge" of the orosensory correlates of the features. This "knowledge" could be used directly as suggested by the bite block results or indirectly in the establishment and maintenance of central programs.

2) As suggested by the discussion in this paper, orosensory feedback might be important in informing any central programming mechanism about the overall state of the system or frame of reference. The use of feedback to make adjustments for changes in the frame of reference could cover a time span corresponding to several movements in a sequence (Polit and Bizzi, 1978). In more general terms, perceptual feedback "regarding the position or movement trajectories of one or more articulators could be used for preprogramming movements several hundred milliseconds into the future." (Larson, personal communication).

3) This paper has implied that central programming plays a major role in the production of articulatory movements. Much of the experimental and theoretical work on other forms of movement suggests that central programming along with internal feedback (feedback entirely internal to the central nervous system) is used for the moment-to-moment (context-dependent) programming of rapid movement sequences. While this is most likely the case for "learned" or "skilled" motor behavior such as speech production (cf. Lindblom, et al., in press), we must keep in mind that vocal tract motor control mechanisms may conceivably have capabilities that other systems do not have (cf. Folkins and Abbs, 1975, 1976, McClean, et al., in press). Thus it is possible that orosensory (peripheral) feedback from the vocal tract is used on a moment-to-moment basis to assist in the programming of articulatory movements in ways that have not been demonstrated for other types of movement.

Ideas such as these are closely related to questions about the nature of fundamental units which underlie the programming of speech production (cf. MacNeilage, 1970). Thus, the difficulty of testing such hypotheses should not stand in the way of exploring them further. The striking similarities between the compensatory articulation and arm movement results discussed above suggest that we may learn increasingly more about speech by continuing to follow future work on analogous types of movement.[3]

References

Abbs, J.H. and G.R. Eilenberg (1976): "Peripheral mechanisms of speech motor control", in Contemporary Issues of Experimental Phonetics, 139-168, N.J. Lass (ed.), New York: Academic Press.

---

(3) I am very grateful to Profs. Emilio Bizzi of M.I.T. and Stephanie Shattuck-Hufnagel of Cornell University for their comments on parts of this manuscript. I also thank Dr. Charles Larson of the University of Washington for his very helpful comments on a previous manuscript. This work was supported by National Institutes of Health Grant NS04332.

Folkins, J.W. and J.H. Abbs (1975): "Lip and jaw motor control during speech: Responses to resistive loading of the jaw", JSHR 18, 1, 207.

Folkins, J.W. and J.H. Abbs (1976): "Additional observations on responses to resistive loading of the jaw", JSHR 19, 820-821.

Leanderson, R. and A. Persson (1972): "The effect of trigeminal nerve block on the articulatory EMG activity of facial muscles", Acta-Oto-Laryngologica 74, 271-278.

Lindblom, B.E.F. (1971): "Neurophysiological representation of speech sounds", Publication 7, Papers from the Institute of Linguistics, University of Stockholm.

Lindblom, B., R. McAllister, and J. Lubker (1977): "Compensatory articulation and the modeling of normal speech production behavior", Paper presented at the Symposium on Articulatory Modeling, Grenoble, France, July 11-12.

Lindblom, B., J. Lubker and, T. Gay (in press): "Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation", JPh.

MacNeilage, P.F. (1970): "The motor control of serial ordering in speech", Psychological Review 77, 182-196.

McClean, M.D., J.W. Folkins, and C.W. Larson (in press): "The role of the perioral reflex in lip motor control for speech", Brain and Language.

Netsell, R., R. Kent, and J. Abbs (1978): "Adjustments of the tongue and lips to fixed jaw positions during speech: a preliminary report", Paper presented at the Conference on Speech Motor Control, University of Wisconsin, Madison, June 2-3.

Perkell, J.S. (1979a): "Phonetic features and the physiology of speech production", in Language Production, B. Butterworth (ed.), New York: Academic Press.

Perkell, J.S. (1979b): "On the nature of distinctive features: implications of a preliminary vowel production study", in Frontiers of Speech Communication Research, B. Lindblom and S. Öhman (eds.), London: Academic Press.

Polit, A. and E. Bizzi (1978): "Processes controlling arm movements in monkeys", Science 201, 1235.

Stevens, K.N. and J.S. Perkell (1977): "Speech physiology and phonetic features", in Dynamic Aspects of Speech Production, 323-341, M. Sawashima and F.S. Cooper (eds.), University of Tokyo Press.