

## S p e e c h P r o d u c t i o n

Report: PETER F. MACNEILAGE  
 Co-Report: PETER LADEFOGED  
 Articulatory parameters  
 Co-Report: MASAYUKI SAWASHIMA  
 A supplementary report on speech production

11

41

49

## SPEECH PRODUCTION

Peter F. MacNeilage, University of Texas at Austin, Austin, Texas, USA

To borrow and adapt a phrase from the German psychologist Ebbinghaus (cited by Boring, 1950), the study of speech production has a long past but a short history. Interest in the speech production process was well developed as early as the time of Panini (Allen, 1953). The well established discipline of Articulatory Phonetics has been ostensibly, solely concerned with the production of speech, even though, according to G.O. Russell (cited by Ladefoged, 1975), phoneticians have been "thinking in terms of acoustic fact and using physiological fantasy to express the idea." But it has only been in the last 20-25 years that the scientific study of speech production, using sophisticated instrumental techniques, and the experimental method, has gained any momentum. Now, in 1979, we look at a flourishing discipline. My task is to convey the flavor of this discipline to a wide range of readers. Unfortunately, my own linguistic limitations prevent an adequate coverage of work not written in the English language. Different limitations dictate a neglect of various subtopics, hopefully to be corrected by my co-reporters. These subtopics include supra-segmentals, tone, timing, phonetic influences on sound patterns of languages, and many aspects of speech pathology. The status report falls into three sections: 1. Functional properties of the speech production apparatus; 2. Control principles underlying speech production; and 3. The biological basis of the speech production process.

1. Functional Propertiesa. Respiratory Function

The main function of the respiratory system during speech is to provide a relatively constant level of subglottal pressure which serves as the power source for the speech act. In long stretches of speech following deep inspiration, this is achieved by active muscular forces first combating and then complementing passive forces towards expiration (relaxation pressure) when lung volume is larger than its resting levels, and then combating passive forces towards inspiration when lung volume is smaller than its resting level. For many years now the work of Ladefoged and his

The reports will also be published in Language and Speech, 1980.

colleagues has provided the standard view of this process (e.g. Ladefoged, 1967). According to this view the following sequence of muscular events occurs: first inspiration is accomplished by the combined action of the diaphragm and the external intercostal muscles. In the initial stage of expiration, the external intercostal muscles combat the (expiratory) relaxation pressure. At the point when relaxation pressure alone becomes insufficient to maintain the required subglottal pressure, the internal intercostals begin to exert a gradually increasing expiratory effect. When lung volume becomes less than that at the end of normal expiration, other muscles including abdominal muscles begin to supplement the expiratory effects of the internal intercostals. Normal conversational speech involves a much more restricted range of lung volumes, and only the internal intercostal muscles are required for expiratory control.

Recent work by Hixon et al. (1976) appears to require one major modification of this view. These researchers consider the abdominal musculature to be continuously active under speech conditions, not only during expiration (for which they are anatomically suited) but during inspiration in conversational speech as well.

Hixon et al. consider that the role of abdominal muscle activity during expiration is to allow for more efficient alveolar pressure generation by the rib cage. This effect is explained by analogy with maneuvers that can be carried out with an elongated balloon in which the portion nearest the neck is analogous to the rib cage and the distal portion is analogous to the abdomen. If one squeezes the half of the balloon near the neck manually, to simulate rib cage maneuvers, pressure will build up within the balloon, and simultaneously cause the distal half of the balloon to expand outward. Combating this outward expansion by contracting the abdominal muscles allows a more efficient pressure build up immediately below the neck of the balloon when it is squeezed in that region.

The role of the abdominal muscles in inspiration is considered to be facilitation of the role of the diaphragm. It is noted that in comparison with quiet breathing, speech breathing consists of "extremely abrupt inspirations and considerably prolonged expirations and that short inspiratory periods are desirable for communication purposes". As Hixon et al. put it: "Because of

inward displacement of the abdominal wall, the diaphragm is displaced axially headward such that its principal muscular fibers (costal) become substantially elongated and its radius of curvature increased. The significance of this externally imposed adjustment is that the diaphragm is in effect "mechanically tuned" to a configuration that tends to optimize its potential for producing rapid and forceful inspiratory efforts."

In addition to providing a relatively constant subglottal pressure level for speech, the respiratory system provides transient increases in subglottal pressure for various suprasegmental and segmental purposes. The precise scope of this second role of the system is not yet well defined partly because of considerable methodological difficulties. Appropriate EMG data is hard to obtain. Body plethysmographs are limited in their sensitivity, and effects on subglottal pressure produced by changes in glottal resistance must be distinguished from effects due to activity of the respiratory system (Ohala, 1974).

#### b. Laryngeal Function

Laryngeal Mechanisms were dealt with quite comprehensively at the last International Congress (Fant and Scully, 1977) and there has not been a great deal of change in our knowledge about them since that time.

In the past few years there has been an increasing realization of the versatility of the vocal folds in producing vocal sound "at a wide range of fundamental frequencies, with great varieties of tonal qualities." (Hirano, 1977). The Myoelastic Aerodynamic Theory of van den Berg (1958) according to which the vocal folds are forced open by increasing subglottal pressure and close again as a result of their own elasticity and the Bernoulli force, remains an appropriate view of the phonation process. But in order to account for the wide range of conditions under which the vocal folds vibrate it has become useful to assume that each vocal fold consists not of a single mass but of a lower and an upper mass roughly corresponding to Hirano's dichotomy between the muscular "Body" of the folds and a mainly ligamentous "Cover", respectively. These two masses move to some extent independently during normal chest register phonation, partly because contraction of the vocalis muscle within each fold sufficiently counteracts the longitudinal tension effect of the cricothyroid to allow the cover to be "loose" and free to vibrate (Hirano, 1977).

We remain relatively uninformed as to how the vibratory patterns of the more unusual but nevertheless linguistically important modes of phonation such as a creaky voice and breathy voice are achieved. With respect to pitch control, a very straightforward relation seems to exist between pitch increase and activity of the cricothyroid muscle (e.g. Atkinson, 1978). On the other hand, there is yet little agreement on how pitch is lowered (Fujimura, 1977a). For one thing, the relative role of the passive effects of reduction in contraction of muscles associated with pitch raising, and the active effects of contraction of pitch lowering muscles situated extrinsic to the larynx has not been satisfactorily established.

Lisker and Abramson's (1971) contention that "the universally most important mechanism for the voiced-voiceless distinction is along the glottal adduction-abduction dimension" (Fant, 1977) is widely accepted. As Fant notes: "the posterior cricoarytenoid muscle... which is the only abductor would accordingly be responsible for glottal opening and thus devoicing in consonants irrespective of the degree of aspiration" (Fant, 1977). The interarytenoid muscle plays the main role in adduction.

The functional role of larynx height in the achievement of both the voiced-voiceless distinction, and the control of pitch is not yet understood. Larynx elevation is positively correlated with both devoicing, and pitch increase, but why this is the case has yet to be explained.

Of the three components of the speech production apparatus the laryngeal component has benefitted from the most sophisticated modelling of the interaction of aerodynamic and biomechanical influences. Prominent examples of recent models of vocal fold vibration are those of Flanagan et al. (1975) and Titze (1976).

#### c. Articulatory Function

The articulatory system is by far the most complex of the three components of the speech production apparatus. A great deal is now known about the way in which vocal tract area functions (shapes) serve to modulate the glottal sound source for speech. But in recent years our knowledge of vocal tract shapes has been pushed beyond the characterizations of traditional articulatory phonetics in two important ways. First, the postures actually adopted by the articulators have become better understood.

Second, using both traditional and new experimental techniques we have gained a good deal of new information about articulatory dynamics (Sawashima and Cooper, 1977). The following is a brief review of some examples of recent developments, intended to illustrate the diversity of motivation and method characteristic of this area of interest.

Some progress is signified by our greater readiness to accept the fact that speech is an "output oriented" activity (Fant, 1977, p. 8). Its aim is to produce an acoustic signal adequate to convey a linguistic message. Because of the non-uniqueness of the relation between vocal tract shapes and acoustic waveforms different speakers are able to communicate the same message with different articulatory postures. From the study of X-ray movies of 5 speakers, Ladefoged et al. (1972) showed that the traditional characterization of vowels in terms of the high-low and front-back dimensions of the tongue is not appropriate, and that there is considerable variation in the tongue configurations adopted by different speakers producing the same vowel. In a similar vein, Bell Berti (1975) has described individual differences in articulatory maneuvers assisting in control of the voiced-voiceless distinction, primarily by controlling vocal cavity volume so as to influence the pressure drop across the glottis.

The articulatory system consists of a set of interdependent structures innervated by a large number of muscles. Part of the search for functional principles underlying articulation has been an attempt to define the number of degrees of freedom in the operation of the system. For example Ladefoged and his colleagues (Harshman et al., 1977; Ladefoged, 1977) have used the statistical technique of Factor Analysis in an attempt to define the number of degrees of freedom in the production of tongue shapes for English vowels. Their analysis revealed two components, one representing "an upward and backward movement of the tongue", and the other representing "a forward movement of the tongue together with a raising of the front of the tongue" (Ladefoged, 1977, p. 217). Ladefoged notes that the former component can be thought of in terms of the action of the styloglossus muscle and the latter in terms of the action of the genioglossus muscle. But he cautions that the two components "if they have any physiological reality at all, are best thought of as high level cortical control functions." (p. 218).

A number of researchers have formulated articulatory models in an attempt to characterize various functional aspects of the articulatory system. Lindblom and Sundberg (1971a) have presented a model which is an attempt at an explicit quantitative specification of the contribution of the individual articulatory structures -- the lips, jaw, tongue and larynx (height) -- in the production of vowels. They also consider that tongue positions can be specified by 2 components, choosing the anterior-posterior location of the tongue body, and the extent to which the tongue body has been deformed from its natural shape. They justify the introduction of the jaw as a parameter in articulatory models on the grounds that it "makes it possible to explain why openness occurs as a universal phonetic feature of vowel production." In their view "the degree of opening of a vowel corresponds to a position of the jaw that is optimized in the sense that it cooperates with the tongue in producing the desired area function" (p. 1166).

An approach to modelling the physical properties of the tongue by computer simulation has been reported by Fujimura. The model "consists of 44 tetrahedral elements as internally uniform subunits of a linear elastic medium. These subunits are organized into 14 prism-shaped functional units representing independently controllable substructures." (Fujimura, 1977b, p. 226-7). The input forces, representing both intrinsic and extrinsic lingual muscles, "...can be specified as a linear combination of any number of internally uniformly distributed stresses within specified functional units, and forces acting directly (externally) on any of the nodal points of these units" (Fujimura and Kakita, 1978). The choice of forces is guided by EMG studies of the activity of lingual muscles during various speech gestures. An example of the outcome of this work, which is still in its earlier stages, is the interesting claim that the required vocal tract configuration for /i/ is relatively insensitive to the precise amount of contraction of the genioglossus muscle. This claim is analogous, at the physiological level, to the Quantal Theory of Stevens, based on observations of the relation between articulatory configurations and sound attributes: "For a particular range of an articulatory parameter, the acoustic output from the vocal tract seems to have a distinctive attribute that is significantly different from the acoustic attributes for some other region of the articulatory

parameter. Within this range of articulation, the acoustic attribute is relatively insensitive to perturbations in the position of the relevant articulatory structure." (Stevens and Perkell, 1977, p. 324).

If these approaches to articulation have anything in common, it is an attempt to define the constraints that determine the observed articulatory events and the absence of other articulatory events that seem logically possible. Attempts have also been made to specify constraints on articulatory dynamics associated with the production of stress and with changes in speaking rate. In an initial spectrographic approach to these questions Lindblom (1963, 1964) concluded that, in Swedish, vowel reduction in unstressed syllables and at faster speaking rates might simply be a mechanical result of the decreased time available for articulatory movements under these conditions. More recent EMG studies (e.g. Gay, 1977) have shown instead that there are differences in control signals to the articulators when stress level or speaking rate is changed. Apparently these stress and rate dependent changes in control cannot be accounted for in terms of any one simple algorithm. Consonants and vowels must be considered separately, as reduction effects are greater in vowels and segment durations reduce more in vowels than in consonants. Stress and rate effects are not always the same. Whereas vowel reduction is characteristic of unstressed syllables, it is only one of the 2 choices of an individual speaker in increasing speaking rate, the other being an increased rate of articulator movement to avoid reduction (MacNeilage, 1978a). Even reduction, when it occurs, is not simply accomplished by a uniform reduction in force of articulation. Amount of undershoot has been observed to differ on different vowels (Gay, 1977). The intuition that stress and rate modifications can be achieved by merely changing the values of some general time-dependent motor control variable has not yet been adequately supported.

## 2. Control Principles

Parallel with work on the functional properties of the speech apparatus has been a concern with the general control principles underlying speech production. Interest has focussed on easily identifiable articulatory gestures associated with individual speech segments -- particular tongue and lip configurations, jaw

positions and velar positions. A most deep-seated conviction is that there must be some invariance underlying the achievement of a configuration for a particular vowel or consonant, regardless of its segmental context. An early hypothesis was that this invariance might lie in the motor command sent to the muscles and observable by means of electromyograms (EMG). However, EMG studies showed, on the contrary, that context dependence in motor commands was the rule (MacNeilage, 1970). For other approaches I quote extensively from a recent review (MacNeilage, 1978a):

"Another group of theorists focussed on the fact that the results of gestures associated with a given phoneme (i.e. the positions achieved by them) remained relatively invariant in different contexts and suggested that therefore gestures were controlled in terms of the specification of invariant goals or targets. As to the nature of these goals or targets, I suggested in 1970 that they could be points specified within an internalized space co-ordinate system of the kind Lashley (1951) considered to underlie all movement control (MacNeilage, 1970). One indirect argument for this view is that visual-motor coordination is certainly guided by an abstract conception of space and therefore the auditory-motor coordination of speech may be also. In addition, control of the speech apparatus in the absence of an auditory component, as in the acts of mastication, and perception of oral stereognosis, would seem to require an abstract spatial analysis mechanism.

Informal evidence of the controlling role of goal or target specification during speech can be obtained by observing a speaker speaking with clenched teeth. Under this condition, acoustic output seems minimally impaired, suggesting that goals are successfully approximated, even though extensive compensatory articulation is probably required. More formal evidence comes from Lindblom and his colleagues who have twice performed an experiment in which subjects were required to produce vowels with bite blocks up to 25 mm in size between the teeth (Lindblom and Sundberg, 1971b; Lindblom et al., 1978). They found that immediately after bite block insertion, subjects achieved the correct formant frequencies in the first pitch period of the subsequent vowel. A subsequent midsagittal X-ray of these subjects during vowel production with a bite block inserted showed close approximation to normal vocal

tract shapes. This result suggests that even under the bite block condition articulators may be successfully controlled by invariant spatial goals or targets. However, in an experiment in which Folkins and Abbs (1975) unpredictably impeded jaw elevation movements associated with closure for a bilabial stop, the upper lip responded with active compensatory lowering, resulting in bilabial closure at a different (lower) point in space than normally observed. Such a finding suggests that the specification of goals or targets may not be in terms of absolute space in this case, but in terms of some other end such as articulator contact, or intra-oral pressure buildup. In addition, goals specified in terms of pressure would seem to be plausible in the respiratory system, where relatively constant subglottal pressure is preserved during speech, using widely varying muscular forces and lung volumes (Hixon et al., 1976)."

"In recent years a number of writers have emphasized the possible role of auditory targets in speech gesture control (e.g. Nootboom, 1970; Ladefoged et al., 1972). Informal evidence for the necessity of auditory targets in some sense of the term is quite conclusive. The auditory information provided by our language community is the only source of goals for our acquisition of speech production. A given auditory goal is sometimes achieved in a single subject by more than one spatial configuration of the speech apparatus. For example single intervocalic [p] is produced in English with vocal fold abduction (Lisker et al., 1969). But cluster-initial intervocalic [p] as in "upbringing" is produced in some subjects by vocal fold adduction (glottal stop) (Westbury, 1978). Thus it is the auditory goal that remains invariant in this case at the expense of invariance in spatial configurations. Further evidence on the relation between internalized auditory standards and movement control comes from an experiment by Riordan (1977). She reported that if rounding gestures of the lips are mechanically prevented, compensatory larynx lowering occurs, to achieve the lengthening of the vocal tract necessary to produce the formant frequencies of rounded vowels. This result shows that the control mechanism is capable of going beyond shape constancy in achieving auditory constancy.

The kinds of targets discussed so far are static targets. But when I produce the diphthong /au/ there is no evidence that

any static auditory or spatial target is being aimed at. In the period during which formant frequencies are relatively unaffected by preceding and following segments, the second formant for /au/ is in continuous motion. The perceptual importance of the dynamic properties of formant transitions, even for vowels, leads us to believe that some specification of dynamic properties must underlie the talker's production of them. Of course, close specification of the dynamics of speech movements is always made by the talker in an utterance whether it has any obvious perceptual consequences or not. Thus the issue to be raised here, ----, is the relation between static and dynamic aspects of the operation of the system." (MacNeilage, 1978a).

A good deal of work has been done on coarticulation, the study of the temporal scope of particular articulatory gestures and how this changes with segmental context. Coarticulation effects have been of interest because of the hope that the precise temporal scope of these effects would provide us with an understanding of the role of various linguistically defined units (e.g. the phoneme, the distinctive feature, the syllable, the word) in the movement control stages of the speech production process. Coarticulatory effects have been observed for up to 7 segments in the anticipatory (right to left) direction (Benguerel and Cowan, 1974) and in the perseveratory (left-to-right) direction (Ghazelli, 1977). Although they occasionally seem to respect syllable boundaries (Ushijima and Hirose, 1974) and word boundaries (Ghazelli, 1977), more often their temporal scope seems independent of the boundaries of linguistic units. They are sometimes not even blocked by mechanical incompatibility between the coarticulatory gesture and gestures for other segments (Sussman et al., 1973). The only thing that seems to reliably block these effects is the avoidance of production of an "immediate successional impact" -- a change in the acoustic properties of a neighboring segment which would change its message status for the listener (Kent and Minifie, 1977). Thus, all in all, the use of coarticulation to determine the basic properties of the control system has been relatively unsuccessful.

In conclusion, it must be conceded that we still know very little about the issue of invariance in the control of gestures or about the principles underlying coarticulation. What we have done so far is little more than to point to aspects or consequences

of gestures that possess invariance and suggest that the goal of the control system must be to achieve this invariance.

In some sense, what we are seeking is biological equivalents of linguistic units. But the precise relation between linguistic units defined primarily by means of analysis of the message structure of language, and control units, compatible with speech signal characteristics, is extremely hard to define. As Stevens and Perkell (1977) point out "There is little argument among students of speech and language that speech events at one level are organized in terms of segments and features." (p. 323). But the lack of argument may only exist because there has been comparatively little effort to reconcile the message and signal levels of conceptualization. The two researchers who have made the greatest recent effort to characterize speech from the traditional viewpoint of articulatory phonetics, Catford (1977) and Ladefoged (1975, 1978) both warn against assuming any simple relation between signals and message units. Catford has concluded that the attempt to define a finite universally applicable set of distinctive features is at best procrustean. Ladefoged concludes that it is erroneous to assume that a phonological feature can be defined in terms of a single physical scale which can be used for specifying contrasts between and within languages. He argues that: "From an acoustic or physiological point of view most phonological features are cover features definable only in terms of complexes of phonetic parameters." In the absence of a straightforward biologically defined relation between observable signal properties and underlying message units it is not clear what it means to assert that speech events at one level are organized-- "in terms of features".

Speech errors have been an important source of the comparatively rare information which bears on what one can call the psychological reality of linguistic units (Fromkin, 1973). An underlying assumption has been that linguistic units have psychological reality to the extent that they show themselves, in speech errors, to be independently variable in the time domain. Distinctive features do not appear to qualify as psychologically real units on these grounds. In approximately two-thirds of spoonerisms involving segments the target segments (e.g. [l] and [r] in "leaf raking" "reaf laking") differ by only 1 distinctive feature

(MacKay, 1970) and thus it is not possible to decide whether a feature or a segment has been exchanged. But where target segments differ by more than one feature (e.g. [p] and [n] in "pointed nail") there is almost never an exchange of a single feature (e.g. "tainted mail") even though, in terms of a feature model, that would be the simplest error. Shattuck-Hufnagel and Klatt (1978) have recently concluded that "features are not independently movable entities at the level where most substitution and exchange errors are made." Neither do syllables move around in speech errors. Nor do speech error data encourage the choice of the phoneme as an underlying form at the segmental level because segmental permutations are so restricted by phonotactic factors -- prevocalic, vocalic, and post-vocalic segments exchange with like components -- whereas the phoneme is ideally a context-free entity. Perhaps the most basic unit at the level at which most temporal sequencing errors can occur is the phonotactically restricted allophone.

Speech errors have also been used to try to determine the number of relatively separate stages or levels in the speech production process and the operations which occur at those levels. It remains possible, though difficult to demonstrate that units such as the feature, phoneme, and syllable have psychological reality at levels where operations do not lend themselves to independent variation of those units in speech errors. (For example it has been suggested (Fry, 1964) that the syllable is a rhythmic entity.) We are still a long way from being able to place sufficient constraints on multistage schemata for speech production. Two illustrations of the problems for such schemata can be given by considering the relation between direct movement control and more underlying levels. On the one hand it has been observed, as mentioned earlier (Westbury, 1978) that the opposite motor control gestures of vocal fold abduction and adduction can be used to achieve examples of what is assumed to be the same underlying voiceless consonant. On the other hand similar motor control gestures can be used to achieve examples of opposing underlying forms. The timing of voice onset and closure release is similar for representation of underlying /d/ in "duck" and (presumably) underlying /t/ in "stuck".

A number of researchers have shown some impatience with efforts to determine the nature of speech control at levels so far

away from direct observation, at the present stage of our knowledge (e.g. Moll et al., 1977). The following paraphrase of the views of Netsell (Moll, 1977) is representative of the concerns of these researchers:

"What existing articulatory data allow us to differentiate the nature of the input commands as phonemes, phones, syllables, words, etc.? Given present methodologies and conceptualizations, what would be the nature of an experiment or experiments that would clarify the character of these motor commands? In relation to these questions, it was noted that inferences from articulatory data can result in varying and contradictory conclusions concerning the nature of the input commands. This results from the fact that articulatory measurements reflect both the effects of the input commands, presumably at a high neural level, and the properties of the intervening physiological systems. Thus, it will not be possible to make inferences about the input command structure until we can separate out the effects of the system characteristics. In addition, it was noted that we must be able to formulate our hypotheses about command units in unambiguous and physiological terms before we can test them effectively by physiological observation." (p. 407).

Netsell's viewpoint emphasizes the importance of understanding the peripheral neuromechanical properties of the speech production system, and from this perspective the relative role of feedback or closed loop control becomes a central issue. On this issue most would agree with Stevens' (1977) view that: "Since speech can be considered to be a habituated or stereotyped form of motor behavior (which does not usually encounter external disturbances), preplanning mechanisms are probably used much more heavily than peripheral feedback for the moment-to-moment control of vocal-tract movements." (p. 343-346). However, Abbs and his associates have pointed out that: "regardless of the nature of the underlying neuromuscular mechanism controlling speech production, evaluation of the system biomechanics is necessary" (Muller et al., 1977). They note that: "if an open loop mode of control is hypothesized then we must concede that the system is either quite knowledgeable about the biomechanics and compensates for them during the control process of organizing the movement, or that the biomechanical characteristics are relatively simple and need not

be considered by the central mechanism." Up until the last few years, most conceptions of the upper articulators assumed a third possibility. Mechanical properties were considered to limit a speaker's ability to produce an invariant output for a phonological unit, resulting in undershoot and coarticulatory effects which the communication system had to tolerate because of neural limitations in production control. (See MacNeilage, 1970, for a review of this view.) But as Abbs and Eilenberg (1976) observed: "it cannot be assumed that peripheral mechanical influences are limiting in their influence upon speech movements. The passive properties of inertia and elasticity most appropriately are considered energy storage mechanisms, and although they may absorb energy generated during one interval of time, they have the capability to release that energy for later contributions to the system's output." (p. 142).

It is now well known that the control system has developed the ability to take stored elastic energy into account in the respiratory system. This was evident in the account of respiratory control given earlier. But the relative role of open and closed loop control in the ongoing control of these and other aspects of speech production is not yet well understood. Part of the motivation behind the work of Abbs and his associates on this topic is the assumption that by systems analysis it can be determined whether it is even analytically possible for any articulator movement to be under closed loop control given the transfer function of the particular neuromechanical system under observation, and the observed movement dynamics (Muller et al., 1977; Abbs et al., 1977). Variables involved in this transfer function include mechanical properties of the musculature and its accompanying load, sensory receptor properties, and potential neural transmission circuits with their loop gains. Mechanical and transmission delays in any system are detrimental, because they create phase lags between the system error and the feedback correction signal. Abbs et al. (1977) state that "... it is possible to determine the bandwidth over which a particular afferent loop might contribute; depending on the total phase lag introduced by its components. That is, if a feedback loop is to contribute to the control of movement, it must have a positive gain in the same frequency range as the movement itself."

In recent years there has been a good deal of disillusionment with the main technique used to investigate closed loop con-

trol of speech -- the sensory nerve block technique -- because of the lack of specificity inherent in its application and the possibility that it involves direct motor effects (Borden et al., 1973). Using the more acceptable technique of on-line intervention, Folkins and Abbs have unequivocally shown the operation of closed loop control of jaw elevation by demonstrating a compensatory response to interference with elevation which had a latency of 30 ms. But as Abbs et al. have pointed out, a general answer to the question of the relative role of closed loop control during speech must await a detailed analysis of all the individual system components in the entire speech apparatus.

Another approach to the understanding of peripheral speech mechanisms, favored in our laboratory, is the analysis of the final form of the neural control signals, at the level of the individual motor unit. The motor unit (an individual motoneuron and the muscle fibres that it innervates) has been described as the quantal element of movement control. The entire control signal for speech production can be described in terms of two variables of motor unit function, the number of motor units activated, and the discharge rate of each. We have concerned ourselves with the understanding of these two variables in speech musculature. In addition to providing parametric information which should aid in modelling of the neuromuscular stage of speech, we have determined that (with minor qualifications) the Size Principle (Henneman et al., 1965) is operative in speech control as in many other aspects of human and animal motor control (MacNeilage, 1978b). "Size" here refers to a number of correlated properties. For example: large motor units have larger cell bodies and larger axons, and their axons innervate more and larger muscle fibers. A number of functional properties are considered to relate to cell body size. Larger cell bodies are considered less excitable than small ones and are thus recruited into a movement at higher input levels (that is, later) than small ones. But once activated, larger cell bodies have greater sensitivity to input changes than small ones. They can be said to have higher gains. Larger cell bodies also have shorter afterhyperpolarization (AHP) durations than small ones. For our purposes AHP duration can be regarded as an index of a recovery cycle following a motoneuron discharge (a single firing) so that short AHP durations allow higher discharge rates. This information



can be regarded as a contribution towards defining the terms in which any control decisions are communicated to the speech apparatus.

As a footnote to this section on control of speech production, I give notice that an alternative view to conventional theories of speech production has recently been advanced. Again I quote extensively from a recent review (MacNeilage, 1978a): "The view arises from what has come to be known as Action Theory which has as its aim a general theory of coordinated movement. (Perhaps a better term at present would be Action Metatheory as it consists primarily of ideas about the form that theory of action should take.) Action Theory owes its origin primarily to the Russian physiologist Bernstein (1967) and has been developed in this country particularly by Greene (1972) and Turvey (1977) and with respect to speech by Fowler (1977, Fowler et al., 1978). The theory calls for a radical reformulation of the theory of speech production. Current speech production theories, that assume underlying units, and elaborate processes governing the surface manifestation of these units are dubbed "translation theories" (Fowler et al., 1978). These theories are considered to be unnecessary, as the linguistic units are, in some sense, directly and invariantly represented in the output, and do not exist independent of that representation. Temporal and spatial aspects of control are regarded as being integral to each other and therefore not to be considered separately. The two components of current models, the basic segmental specification, and timing schemes are thus also considered integral to each other and timing is described as intrinsic rather than extrinsic. This means that the timing-determined properties of the output arise naturally from its intrinsic organization.

The central concept of action theory is that of the "coordinative structure". A coordinative structure is defined as "a set of muscles, often spanning many joints, that is constrained to act as a unit" (Turvey et al., 1978). These structures are considered to be established by biasings of reflex circuits referred to as "tunings". Some properties of coordinative structures are modulable; "For example rate of walking and gait are modulable properties of the muscle systems that determine walking" (Fowler, 1977, p. 206). The constraints arising from the coordinative structures

are considered to determine directly what effects the modulable properties will have on movement. For example, I assume speaking rate would be a modulable property of the coordinative structures governing speech production, and thus its effects would arise directly from those structures rather than being imposed on them from an external source.

An act such as an utterance "is believed to be governed by functionally embedding (as opposed to temporally concatenating) coordinative structures. Each nesting level delimits a broader equivalence class of movements than the finer grain level nested within... The more coarse-grained nesting levels are established by altering the relationships among smaller coordinative structures, and at the same time they act as constraints on the lower ones." (Fowler et al., 1978, p. 28). For example, in Fowler's (1977) analysis the most coarse-grained coordinative structure embodies an entire utterance and at the other extreme 4 coordinative structures are proposed for vowels.

A good candidate for one of the most coarse-grained coordinative structures could be one that is responsible for maintaining a relatively constant level of subglottal pressure during a single expiratory phase. One coordinative structure for vowels is considered closely analogous to the state underlying the voluntary assumption of a certain fixed joint angle at the elbow, by a human subject, and the maintenance of that angle in the face of various loads (Asatryan and Fel'dman, 1965). The subject is considered to adopt an arbitrary "zero-state" of the joint-muscle system, thus creating a system with spring-mass properties. The production of /ε/ is considered to be achieved partly by the imposition of a "zero-state of the extrinsic tongue system" (Fowler et al., 1978, p. 71) which will produce tongue elevation following /æ/, but tongue depression following /i/.

It is obviously not possible to do justice to this new theoretical orientation in the space available here. The theory is provocative in its attempt to place speech production in a general biological perspective, and in its implication that speech production is by no means special. Nevertheless it is my overall impression that the relation between traditional theories of speech production and Action Theory has so far only been loosely defined, and that the value of the analogies made between speech and other

coordinated movement sequences still needs to be carefully scrutinized. The specific consequences of Action Theory for speech production have not yet been well established. For example, are changes in vowels with stress and speaking rate consistent in form with some specific coordinative structures with specific modulable properties? Is what is descriptively labelled as "undershoot", a result of changes in the coordinative structures or in the effects of modulable properties? How does the theory handle the case mentioned earlier of two opposite movement outcomes (abduction and adduction) in the achievement of voicelessness for /p/? Do short term memory constraints influence the operation of coordinative structures? It is to be hoped that these and many other problems will be fruitfully addressed as the implications of Action Theory for speech become clearer, and the interface between Action Theory and traditional ideas about speech production control becomes more clearly defined."

### 3. Biological Basis

Naturally occurring language -- including speech -- is still considered species-specific to humans despite the inroads being made by chimpanzees and gorillas learning sign language. But the biological basis for this specificity is not yet established. The considerable linguistic ability displayed by chimpanzees and gorillas using manual signs, together with the failure to teach them language involving speech production, suggests that species-specificity is greater for speech than for language (in the sense of an abstract message system). Species specificity in speech perception ability has not yet been established (though see Warren, 1976). The one-month-old human has shown a spectacular ability to discriminate between linguistically distinctive stimuli, most marked for the voice onset time (VOT) dimension (Eilers, 1978). But this ability, though apparently innate, cannot necessarily be considered either species specific or speech specific, in the light of the perceptual ability displayed by chinchillas with the VOT continuum, and the results of perceptual studies of nonspeech analogs of the VOT continuum (MacNeilage, 1977) including a study of infants (Jusczyk et al., 1977).

With respect to speech production, Lieberman (1975) has made a case that a "crucial" evolutionary development separating modern humans from lower hominids, is the development of the "bent two-

tube" supralaryngeal vocal tract from an earlier single tube configuration. This structural development increases the possible number of vocalic items in a speech sound inventory and makes possible production of the 3 "quantal" vowels /i/, /u/ and /a/. Lieberman (1972) has claimed that: "Modern speakers, in all likelihood make use of these extreme vowels to ascertain the size of the vocal tracts of individual speakers. This information is essential for the speech "decoding" that is the basis of the rapid rate of information transfer of human speech. Neanderthal man, though he could produce part of the human phonetic repertoire, would be incapable of speaking any human language" (p. 272). However, more recently, Verbrugge et al. (1976) have concluded that with respect to perception, "There is little evidence to support a claim of a special role for the point (quantal) vowels." (p. 198). Perhaps the structural development reported by Lieberman has only quantitative significance, making more distinctive sounds possible, though it should be noted that some human languages have quite small segmental inventories (e.g. 15 in New Zealand Maori (Biggs, 1961)).

On the topic of functional aspects of the biological basis of speech production, it is my contention that we have seriously underrated the importance of the "long and for the most part orderly series of stages of prelinguistic vocalization that are quite stereotyped in nature, and occur in the absence of model vocalizations of others" (MacNeilage, 1978a). A case can be made for the innateness and species specificity of many of these vocalizations including babbling (chimpanzees are considered to babble only in a quite restricted sense (Kortlandt, 1973)), and despite the methodological problems attendant on their study, they would seem to be a valuable source of information about the biological basis of speech. The neglect of prelinguistic vocalizations is, in my view, largely the result of the influence of Jakobson's theory of language acquisition (Jakobson, 1968), to which babbling was irrelevant. My view that prelinguistic vocalizations are important comes from recent studies by Oller et al. (1976) and others whom they cite, which show an extremely close relation between the sounds and sound sequences of babbling, even the earliest babbling (6-8 months) and the sounds and sequences used in the first words.

As to sounds, aspirated stops, fricatives and liquids are rare in babbling and in first words while unaspirated stops and glides are frequent in both. As to sequences (or phonotactic constraints), consonant clusters and final consonants (especially voiced final consonants) are rare in babbling and in first words, while initial consonants, especially stops, are common in both. Oller et al. conclude: "after examining our data on babbling it is possible to predict quite accurately the nature of the most commonly reported substitutions and deletions which occur in meaningful child speech." (1976, p. 9). I would go further and guess that it is possible to predict from the babbling of an individual child a great deal about what his/her first words will be like, including initial sounds and sequences, relative preference for reduplicated syllables, and aspects of temporal control such as segment durations and variability in those durations.

From this viewpoint the child's first words can be seen as, at least partially, a matter of choosing from the babbling repertoire a set of approximations to adult word forms. As the babbling forms are quite limited, the first words represent an enormous simplification of adult forms, and subsequent learning can be seen primarily as the relaxation of constraints on earlier articulated forms as ability to produce additional forms increases. This view can be contrasted with Jakobson's view of first word learning as a matter of the unfolding of a fixed universal sequence of sound contrasts. (For a review of evidence against this view, see Ferguson, 1978.) The matter of choice of the child's first words was considered by Jakobson to be the result of "a selection by which they become speech sounds only insofar as they are related to language in the strict sense of the word. The selection is therefore inseparably linked to the sign nature of language, that is, is a purely linguistic matter." In contrast: "The question of the prelanguage babbling period proves to be, on the contrary, one of external phonetics, predominantly articulatory in nature ..." (1968, p. 27). A problem of Jakobson's viewpoint is that an appeal to the sign function of sounds gives no explanation of the particular order in which sounds do tend to appear in the first words. (The notion of maximal sound contrast has quite limited explanatory power in this respect.) But if one considers speech as an extension of babbling and it is seen that the first sounds of speech

tend to be the first sounds of babbling then a good deal of the order in sound acquisition becomes at least potentially explicable in terms of the biology of the production mechanism. One thus loses the need for innate components of speech such as a sign function. The sign function of language may evolve in the framework of the biology of the signalling mechanism and not vice versa. Two examples may illustrate this point. The voiceless unaspirated stop consonant is the only universal stop consonant category and it is the first stop consonant to be observed in babbling and in the child's first words. The Consonant-Vowel syllable is the only universal syllable type, and it is the first to be observed in canonical form (i.e. with the time-space constraints typical of speech) in the babbling stage and in the child's first words. Babbling can be regarded as the functional skeleton on which child speech is built and the functional skeleton on which sound patterns of languages are built.

There are two important qualifications to be made about the view expressed here. First, the tendency towards a fixed order of acquisition of sounds is just a tendency. Numerous individual exceptions have been noted. These exceptions mean that no theory that includes a single fixed order of acquisition of sounds or rules, can be valid. Second, although much emphasis has been placed here on prelinguistic vocalizations being carried over into speech without change in form, the functional significance of the babbling stage does not inhere in these forms alone. Other non-canonical forms such as the popular (universal?) 2-3 second long "raspberry" (Oller, 1978) may serve some purpose. In addition both my children (and others) have favored the production of C-V alternations with lateral manual stop consonants, produced by placing the back of the hand in a horizontal position over the mouth at the proximal finger joints, and alternately flexing and extending the fingers. These behaviors suggest that the babbling stage provides a functional skeleton for speech production in a sense that includes, but also transcends, motor stereotypies produced with the time-space properties and the apparatus later used for speech production.

In summary, I believe a case can be made for the innateness and species-specificity of many aspects of prelinguistic vocalization, which may provide a functional skeleton for the development

of the speech production process, both in terms of specific motor behavior and in terms of more general functional prerequisites.

Another source of evidence that has been widely used to claim an innate species-specific basis for language and speech is that of left hemisphere specialization. However, it is important to note that this evidence cannot be used to support a "speech is special" notion. Studies by Kimura and her associates (Kimura and Archibald, 1974; Mateer and Kimura, 1977) have shown that left hemisphere damage, whether accompanied by aphasia or not, results in impairment of reproduction of sequences of movements, whereas right hemisphere damage does not. These results have led Kimura (1976) to suggest that underlying left hemisphere specialization for language and speech may be specialization for the control of skilled motor sequences. In addition, recent studies show that chimpanzees have an anatomical differentiation between left and right hemispheres analogous to that found in humans (Galaburda et al., 1978). These findings, taken together with the failure to induce vocally based linguistic communication in chimpanzees, encourage us to view left hemispheric specialization for speech in man within a broad biological perspective. It is my belief that a good deal of our present knowledge is consistent with the following proposition: left hemisphere specialization for language and speech is derivative of the specialization of the left hemisphere for the control of skilled voluntary movement sequences, and for the perceptual analysis of stimuli related to the movement control skills.

As to the means by which the left hemisphere controls speech production, most speculation, based primarily on aphasia, has been directed towards the localization of functions in the cerebral cortex. Very little information is available about exactly what these functions are, and how they are controlled. The most acceptable general proposition is that whereas damage to anterior cortex results in a nonfluent aphasia, most typically classifiable as Broca's Aphasia, damage to posterior cortex results in a fluent aphasia typically classifiable as Wernicke's Aphasia. In the well known schema of Geschwind (1972) the speech production deficits of Wernicke's Aphasia are due to damage to the mechanism responsible for auditory patterns controlling output, and Broca's Aphasia deficits are in production of speech movements from the auditory

patterns. An additional syndrome of Conduction Aphasia, (now widely accepted by aphasiologists (Green and Howes, 1977) was considered due to a disconnection between the auditory and motor centers resulting from a lesion of the arcuate fasciculus, in the parietal lobe. In the light of the probable role of somesthetic functions, including spatial conceptualization in speech production control, and in light of the parietal lesion site for Conduction Aphasia, I have suggested (MacNeilage, 1978b) that Geschwind's explanation of Conduction Aphasia be re-examined and perhaps supplemented with another. In this view, Conduction Aphasia could at least in part be regarded as a deficit in somesthetic conceptualization underlying spatial target assignment for speech movement control. There were 4 reasons for this suggestion which are discussed in detail elsewhere (MacNeilage, 1978b): "First, the dominant speech production symptom of phonemic paraphasias is consistent with this view. Second, conduction aphasia is typically, though not always associated with apraxia, and apraxia would be expected to result from a deficit in spatial target function. Third, the lesions associated with the syndrome are in the parietal lobe which has traditionally been associated with spatial functions. Fourth, there is no alternative explanation in the literature that has a good claim to being preferable to the one I am suggesting."

Since this view was put forward, apparent counterevidence has come from studies of cerebral blood flow as an index of localization of cortical activity during speech production (Lassen et al., 1978). According to this index, posterior inferior parietal cortex activity was as low as that in several other sites, not expected to be involved in speech production (e.g. occipital cortex). As expected, high levels of activity were observed in Wernicke's Area and in primary sensorimotor cortex serving the speech apparatus (pre- and post-central gyri) and some elevation of activity was observed in Broca's Area. In addition, a surprising finding was the high levels of activity observed in a relatively large region of Supplementary Motor Cortex during speech as well as other motor functions. This region, in superior medial frontal cortex was implicated in speech control by the work of Penfield and his associates (Penfield and Roberts, 1959) but has received little attention since that time. Lassen and his associates con-

sider that the Supplementary Motor Cortex "is involved in the planning of sequential motor tasks" (p. 69). But although the findings of Penfield and Roberts corroborate the blood flow studies with respect to the involvement of Supplementary Motor Cortex (and Wernicke's and Broca's areas), they did not concur in observing a lack of involvement of posterior inferior parietal cortex. Evidence from both electrical stimulation and excision strongly implicated the parietal area in speech control. A resolution of this dilemma apparently awaits a better understanding of the meaning of the evidence from these 3 different sources.

Another finding of the work on cerebral blood flow was that of high levels of activity in the right hemisphere analogs of the left hemisphere regions which showed high activity during speech. The rarity of aphasia from right hemisphere lesions, the absence of right hemisphere initiation of speech in commissurotomy patients and the very limited scope of right hemisphere initiation of speech in left hemispherectomized patients (Searleman, 1977) suggests that right hemisphere does not play a necessary role in speech production. Nevertheless the blood flow studies suggest that the right hemisphere may normally share in speech production to a greater extent than had been previously supposed. The result also suggests that patients with left hemisphere lesions may be able to make more use of the right hemisphere to control speech than most current views of aphasia imply.

#### References

- Abbs, J.H. and G.R. Eilenberg (1976): "Peripheral mechanisms of speech motor control", in Contemporary issues in experimental phonetics, N.J. Lass (ed.), 139-168, New York: Academic Press.
- Abbs, J.H., E.M. Muller, M. Hassul, and R. Netsell (1977): "A systems analysis of possible afferent contributions to lip movement control", paper presented to the American Speech and Hearing Association, Chicago, 1977.
- Allen, W.S. (1953): Phonetics in ancient India, London: Oxford University Press.
- Asatryan, D. and A. Fel'dman (1965): "Functional tuning of the nervous system with control of movement or maintenance of a steady posture -- I. Mechanographic analyses of the work of the joint on execution of a postural task", Biophysics 10, 925-935.
- Atkinson, J.E. (1978): "Correlation analysis of the physiological factors controlling fundamental voice frequency", JASA 63, 211-222.
- Bell-Berti, F. (1975): "Control of pharyngeal cavity size for English voiced and voiceless stops", JASA 57, 456-61.
- Benguerel, A.-P., and H.A. Cowan (1974): "Coarticulation of upper lip protrusion in French", Phonetica 30, 41-55.
- Bernstein, N. (1967): The coordination and regulation of movements, London: Pergamon Press.
- Biggs, E. (1961): "The structure of New Zealand Maaori", Anthropological Linguistics 3, 1-53.
- Borden, G.J., K.S. Harris and L. Catena (1973): "Oral feedback II. An electromyographic study of speech under nerve block anesthesia", JPh 1, 297-308.
- Boring, E.G. (1950): A history of experimental psychology, New York: Appleton Century Crofts.
- Catford, J.C. (1977): Fundamental problems in phonetics, Bloomington: Indiana University Press.
- Eilers, R. (1978): "The complex nature of infant speech perception", Proceedings of the conference on Child Phonology: Perception, Production and Deviation, G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.) (In preparation).
- Fant, G. (1977): "Introductory remarks - Laryngeal mechanisms and features", in "The larynx and language", G. Fant and C. Scully (eds.), Phonetica 34, 252-255.
- Fant, G. and C. Scully (eds.) (1977): "The larynx and language", Phonetica 34, 247-325.
- Ferguson, C.A. (1978): "Learning to pronounce: The earliest stages of phonological development in the child", in Communicative and cognitive abilities - early behavioral assessment, F.D. Minifie and L.L. Lloyd (eds.), Baltimore, Md.: University Park Press (In press).
- Flanagan, J.L., K. Ishizaka, and K.L. Shipley (1975): "Synthesis of speech from a dynamic model of the vocal cords and vocal tract", Bell Syst. Tech. J. 54, 485-506.
- Folkins, J.W. and J.H. Abbs (1975): "Lip and jaw motor control during speech responses to resistive loading of the jaw", JSHR 18, 207-220.
- Fowler, C.A. (1977): Timing control in speech production, Bloomington, Ind.: Indiana University Linguistics Club.
- Fowler, C.A., P. Rubin, R.E. Remez, and M.T. Turvey (1978): "Implications for speech production of a general theory of action", in Language production, B. Butterworth (ed.), New York: Academic Press. (In press).
- Fromkin, V.A. (1973): Speech errors as linguistic evidence, The Hague: Mouton.
- Fry, D.B. (1964): "The functions of the syllable", Zs.f.Ph. 17, 215-237.
- Fujimura, O. (1977a): "Control of the larynx in speech", Phonetica 34, 280-288.
- Fujimura, O. (1977b): "Model studies of tongue gestures and the derivation of vocal tract area functions", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.) 225-232, Tokyo: University of Tokyo Press.

- Fujimura, O. and Y. Kakita (1978): "Remarks on quantitative description of the lingual articulation", in Frontiers of speech communication research, B. Lindblom and S. Ohman (eds.), London: Academic Press.
- Galaburda, A.M., M. LeMay, T.L. Kemper and N. Geschwind (1978): "Right-left asymmetries in the brain", Science 199, 852-856.
- Gay, T. (1977): "Cinefluorographic and electromyographic studies of articulatory organization", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 85-102, Tokyo: University of Tokyo Press.
- Geschwind, N. (1972): "Language and the Brain", Scientific American 226, 76-83.
- Ghazelli, S. (1977): Back consonants and backing coarticulation in Arabic, unpublished Ph.D. dissertation, University of Texas at Austin.
- Green, E. and D.H. Howes (1977): "The nature of conduction aphasia: A study of anatomic and clinical features and of underlying mechanisms", in Studies in Neurolinguistics, Vol. 3, H. and H.A. Whitaker (eds.), 123-156, New York: Academic Press.
- Greene, P.H. (1972): "Problems of organization of motor systems", in Progress in Theoretical Biology, Vol. 2, R. Rosen and F. Snell (eds.), 304-322, New York: Academic Press.
- Harshman, R., P. Ladefoged and L. Goldstein (1977): "Factor analysis of tongue shapes", JASA 62, 693-707.
- Henneman, E., G. Somjen and D.D. Carpenter (1965): "Functional significance of cell size in motoneurons", J. Neurophysiol. 28, 560-580.
- Hirano, M. (1977): "Structure and vibratory behavior of the vocal folds", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 13-27, Tokyo: University of Tokyo Press.
- Hixon, T.J., J. Mead and M.D. Goldman (1976): "Dynamics of the chest wall during speech production: Function of the thorax, rib cage, diaphragm, and abdomen", JSHR 19, 297-356.
- Jakobson, R. (1941): Kindersprache, Aphasie, und allgemeine Lautgesetze, Uppsala: Almqvist and Wiksell. Reprinted as Child language, aphasia, and phonological universals (1968), The Hague: Mouton.
- Jusczyk, P.W., B.S. Rosner, J.E. Cutting, C.F. Foard and L.B. Smith (1977): "Categorical perception of nonspeech sounds by 2-month-old infants", Percept. Psychophys. 21, 50-54.
- Kent, R.D. and F.D. Minifie (1977): "Coarticulation in recent speech production models", JPh 5, 115-134.
- Kimura, D. (1976): "The neural basis of language qua gesture", in Studies in Neurolinguistics, Vol. 2, H. and H.A. Whitaker (eds.), 145-156, New York: Academic Press.
- Kimura, D. and Y. Archibald (1974): "Motor functions of the left hemisphere", Brain 97, 337-350.
- Kortlandt, A. (1973): Discussion of paper by G.W. Hewes, "Primate Communication and the gestural origin of language", Current Anthropology 14, 13-14.

- Ladefoged, P. (1967): Three areas of experimental phonetics, London: Oxford University Press.
- Ladefoged, P. (1975): A course in phonetics, New York: Harcourt Brace Jovanovich.
- Ladefoged, P. (1977): "The description of tongue shapes", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.) 209-222, Tokyo: University of Tokyo Press.
- Ladefoged, P., J. DeClerk, M. Lindau, and G.A. Papçun (1972): "An auditory-motor theory of speech production", U.C.L.A. Working Papers in Phonetics 22, 48-75.
- Lashley, K.S. (1951): "The problem of serial order in behavior", in Cerebral mechanisms in behavior, L.A. Jeffress (ed.), The Hixon Symposium, New York: Wiley.
- Lassen, N.A., D.H. Ingvar, and E. Skinhøj (1978): "Brain function and blood flow", Scientific American 239, 62-71.
- Lieberman, P. (1972): Discussion in Proceedings of the Seventh International Congress of Phonetic Sciences, A. Rigault and R. Charbonneau (eds.), 272, The Hague: Mouton.
- Lieberman, P. (1975): On the origins of language, New York: Macmillan.
- Lindblom, B.E.F. (1963): "Spectrographic study of vowel reduction", JASA 35, 1773-1781.
- Lindblom, B.E.F. (1964): "Articulatory activity in vowels", STL-QPSR, 3.
- Lindblom, B., J. Lubker and T. Gay (1978): "Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation", JPh (in press).
- Lindblom, B.E.F. and J.E.F. Sundberg (1971a): "Acoustical consequences of lip, tongue, jaw and larynx movement", JASA 50, 1166-1179.
- Lindblom, B.E.F. and J. Sundberg (1971b): "Neurophysiological representation of speech sounds". Paper presented at the XVth World Congress of Logopedics and Phoniatrics, Buenos Aires, Argentina, August.
- Lisker, L., A.S. Abramson, F.S. Cooper, and M.H. Schvey (1969): "Transillumination of the glottis in running speech", JASA 45, 1544-1546.
- Lisker, L. and A.S. Abramson (1972): "Glottal modes in consonant distinctions", in Proceedings of the Seventh International Congress of Phonetic Sciences, A. Rigault and R. Charbonneau (eds.), 366-370, The Hague: Mouton.
- MacKay, D.G. (1970): "Spoonerisms: The structure of errors in the serial order of speech", Neuropsychologia 8, 323-350.
- MacNeilage, P.F. (1970): "Motor control of serial ordering of speech", Psych. Rev. 77, 182-196.
- MacNeilage, P.F. (1977): "Is the speaker-hearer a special hearer?", in Modes of perceiving and processing information, H.L. Pick and E. Saltzman (eds.), 53-66, New Jersey: L. Erlbaum Assoc.

- MacNeilage, P.F. (1978a): "The control of speech production", Proceedings of the conference on Child Phonology: Perception, Production and Deviation, G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.). (In preparation).
- MacNeilage, P.F. (1978b): "Neural mechanisms in speech production", in Current issues in the phonetic sciences, H. and P.A. Hollien (eds.), Amsterdam: John Benjamins B.V. (In preparation).
- Mateer, C. and D. Kimura (1977): "Impairment of nonverbal oral movements in aphasia", Brain and Language 4, 262-276.
- Moll, K. (1977): "Physiological studies of speech production: Supralaryngeal", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 404-408, Tokyo: University of Tokyo Press.
- Moll, K.L., G.N. Zimmerman and A. Smith (1977): "The study of speech production as a human neuromotor system", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 107-128, Tokyo: University of Tokyo Press.
- Muller, E., J. Abbs, J. Kennedy and C. Larson (1977): "Significance of perioral biomechanics to lip movements during speech", paper presented to the American Speech and Hearing Association, Chicago.
- Nooteboom, S. (1970): "The target theory of speech production", IPO Annual Progress Report 5, 51-55.
- Ohala, J. (1974): "A mathematical model of speech aerodynamics", Proc. Speech Comm. Sem., Stockholm, 65-72, Uppsala: Almqvist and Wiksell.
- Oller, D.K. (1978): "The emergence of the sounds of speech in infancy", Proceedings of the conference on Child Phonology: Perception, Production and Deviation, G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.). (In preparation).
- Oller, D.K., L.A. Wieman, W.J. Doyle, and C. Ross (1976): "Infant babbling and speech", Journal of Child Language 3, 1-11.
- Penfield, W. and L. Roberts (1959): Speech and brain mechanisms, Princeton, N.J.: Princeton University Press.
- Riordan, C.J. (1977): "Control of vocal-tract length in speech", JASA 62, 998-1002.
- Sawashima, M. and F.S. Cooper (eds.) (1977): Dynamic aspects of speech production, Tokyo: University of Tokyo Press.
- Searleman, A. (1977): "A review of right hemisphere linguistic capabilities", Psych. Bull. 84, 503-528.
- Shattuck-Hufnagel, S.R. and D.H. Klatt (1978): "Symmetry in the direction of substitution for segmental speech errors", JVLVB (in press).
- Stevens, K.N. (1977): Discussion in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 343-344, Tokyo: University of Tokyo Press.
- Stevens, K.N. and J.S. Perkell (1977): "Speech physiology and phonetic features", in Dynamic aspects of speech production, M. Sawashima and F.S. Cooper (eds.), 323-341, Tokyo: University of Tokyo Press.
- Sussman, H.M., P.F. MacNeilage and R.J. Hanson (1973): "Labial and mandibular mechanics during the production of bilabial stop consonants", JSHR 16, 397-420.
- Titze, I.R. (1976): "On the mechanics of vocal-fold vibration", JASA 60, 1366-1380.
- Turvey, M. (1977): "Preliminaries to a theory of action with reference to vision", in Perceiving, acting and knowing: Toward an ecological psychology, R. Shaw and J. Bransford (eds.), Hillsdale, N.J.: Erlbaum.
- Turvey, M.T., R.E. Shaw, and W. Mace (1978): "Issues in the theory of action", in Attention and performance VII, J. Requin (ed.), Hillsdale, N.J.: Erlbaum. (In press).
- Ushijima, T. and H. Hirose (1974): "Electromyographic study of the velum during speech", JPh 2, 315-326.
- van den Berg, J. (1958): "Myoelastic-aerodynamic theory of voice production", JSHR 1, 227-244.
- Verbrugge, R.R., W. Strange, D.P. Shankweiler, and T.R. Edman (1976): "What information enables a listener to map a talker's vowel space?", JASA 60, 198-212.
- Warren, R.M. (1976): "Auditory perception and speech evolution", in Origins and evolution of language and speech, S.R. Harnad, H.D. Steklis and J. Lancaster (eds.), 708-717, New York: New York Academy of Sciences.
- Westbury, J. (1978): Aspects of the temporal control of voicing in consonant clusters in English, unpublished Ph.D. dissertation, University of Texas at Austin.