327

# Use of Feedback in Established and Developing Speech

**GLORIA J. BORDEN**
*Department of Speech*
*Temple University*
*Philadelphia, Pennsylvania*
*and*
*Haskins Laboratories*
*New Haven, Connecticut*

## I. INTRODUCTION

During spontaneous speech, we create utterances in our minds as we produce them out loud. As we formulate the next linguistic chunk to be spoken, we hold it momentarily, if only parts of it, as a perceptual image. We not only know what it is, in general, that we plan to say, but we have a rather abstract idea of how it will sound. Unconsciously, we also have an image of how it will feel in terms of touch, pressure sensations, and movements and positions of speech organs. We know these things because we know our own vocal tract possibilities and our own voicing capabilities. We have heard and felt ourselves talk for years. In almost every speaking situation, we are able to feel and hear ourselves speak. We know so well how we will sound and feel when producing speech that we can continue to produce perfectly intelligible speech in artificial situations

*SPEECH AND LANGUAGE: Advances in Basic Research and Practice, Vol. 3*

ISBN 0-12-608603-6

in which we are prevented from hearing ourselves, as under auditory masking, or are prevented from feeling surface sensations, as under oral anesthesia. It is likely that, in these instances, we continue to receive information on our performance from our muscles and from feedback mechanisms contained within the central nervous system (CNS).

Less skilled speakers must depend upon auditory and tactile feedback more than speakers who have well-established speech-production systems. Young children developing speech, and speakers of any age attempting to learn a new language, must use all available feedback channels in their efforts to match the sound patterns of the new language with the sensations produced by their own imitations. Children with congenital hearing losses are evidently at a serious disadvantage in developing natural-sounding speech patterns. In contrast, those who have learned to speak before the onset of deafness produce good speech with only slight deterioration of intelligibility which takes place gradually after a period of time (Fry, 1966). The use that one is able to make of feedback from one's own speech seems to vary with age as well. Children first acquiring language, whether they are learning one or two languages, are particularly adept at matching their own speech to the models provided. They use feedback to emulate the segmental, intonational, and rhythmic characteristics of the languages to which they are exposed. Too often, however, older people learning a second language fail in their efforts to match the segmental and suprasegmental aspects of the language as spoken, even though they may have mastered the grammar and vocabulary. The auditory, tactile, and muscle-moving images that they have stored for their first language seem to supersede any new images. Thus, adults speak the second language with an accent: the sound patterns of the first language persist in the second.

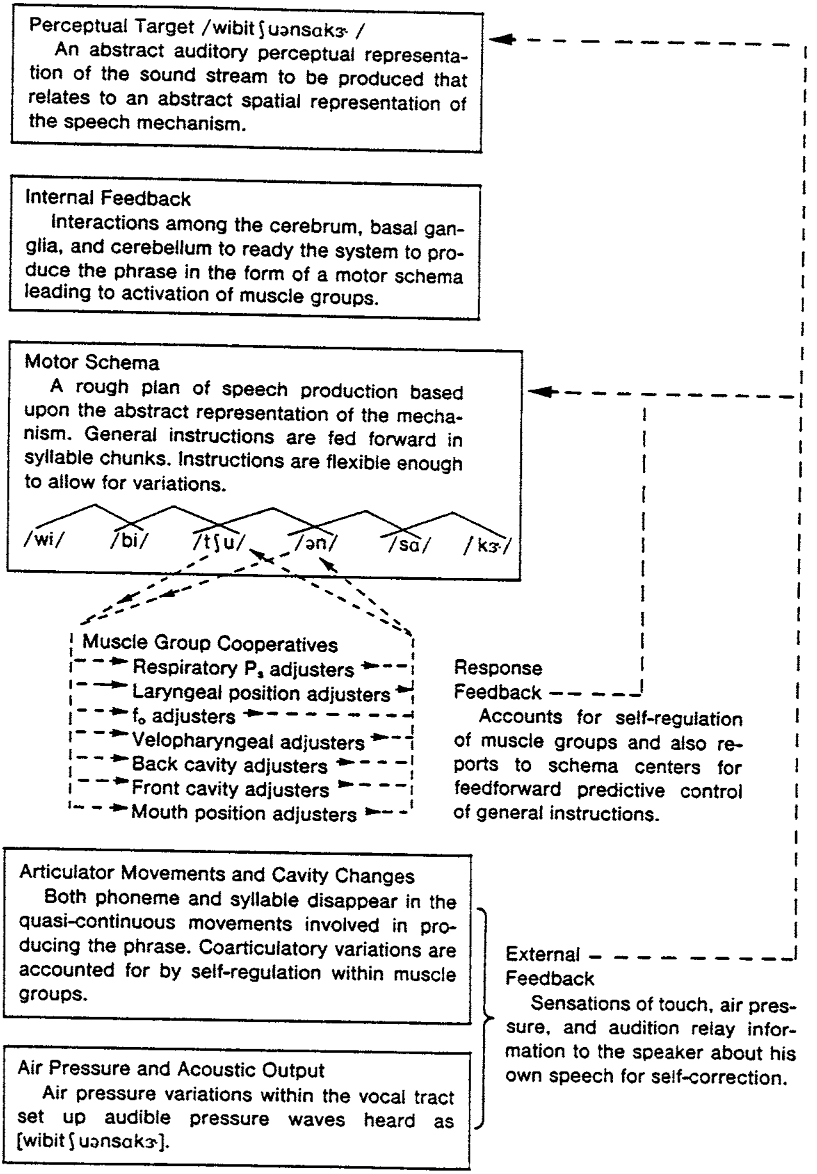
In this chapter, the goal is not to provide a comprehensive review of the literature on motor control, but rather to state a current view of how speech control may operate in skilled and nonskilled speakers—and to include examples from recent research in support of this view. However, before discussing the ways in which feedback may operate during speech acquisition and during established speech, a brief consideration of the control mechanisms themselves and the experimental effects of altering the information that they provide is in order.

## II. CONTROL MECHANISMS FOR SPEECH

Normal speakers obtain feedback on their speech performance at a minimum of three levels of motor organization. Borrowing the terms from Evarts (1971) writing on limb control—and applying them to speech—the levels are (1) internal feedback; (2) response feedback; and (3) knowledge of results or external feedback (Borden, 1979). Internal feedback is a network for information exchange entirely within the brain. The circuit includes the basal ganglia of the midbrain, the motor centers of the cerebrum, and the coordinative centers of the cerebellum. Response feedback is information from the joints, tendons, and especially from muscles, providing position and movement sense. This sensation, often termed "proprioception" after Sherrington (1906) or "kinesthesia," a term that connotes awareness of the proprioception, arises as a response from the motor activity itself. The third level, external feedback, is information based upon the results of the motor patterns, not upon the motor patterns themselves. Knowledge of results as applied to speech would include auditory and tactile information. The air- and bone- conducted sounds of speech are available to the speaker as are sensations of touch and of air-pressure changes (Stevens & Perkell, 1977).

Figure 1 illustrates the multilevel control of speech suggested. In this model of speech production (Borden & Harris, 1980), the speaker is skilled and thus unconsciously knows the general sound of the phrase to be spoken ("We beat you in soccer") as well as the general requirements of the speech mechanism. This knowledge, labeled *Perceptual Target* in the figure, is translated into an abstract motor plan or schema by the interaction of the motor cortex, cerebellum, and basal ganglia of the brain. Neural activity in these areas has been recorded approximately 100 msec before speech; and, although the purpose of the activity is not known, it can be postulated that the perceptual target is being translated into appropriate motor programs. Since the general schema of appropriate motor plans should be well known to the skilled speaker, the cerebellum can cooperate with the midbrain and cerebrum to apply well-practiced sequences of coordinated activities. An internal feedback loop could permit these centers to receive information from one another. The prespeech activities represented in the top three boxes of Fig. 1 are considered to overlap in time, as do those represented in the two boxes at the bottom of the figure. As the motor schema develops, it is implemented sequentially. Kimura (1977) suggests, on the basis of evidence from patients with unilateral cerebral lesions, that the left hemisphere usually dominates in the control of the transitions within sequences of different motor gestures. Thus, the left hemisphere may play a major role in switching from one muscle group to another.

Many groups of muscles cooperate to control the systems important for vocal tract shaping and for sound production. We know that the speech-production system compensates easily for any constraint put upon it—whether the constraint is imposed from without, as in MacNeilage's (1970) example of speaking with a pipe clenched between the teeth, or is imposed by the speech context, as in moving to /t/ from an open vowel rather than a



Perceptual Target /wibitʃuənsɑkɝ /

An abstract auditory perceptual representation of the sound stream to be produced that relates to an abstract spatial representation of the speech mechanism.

Internal Feedback

Interactions among the cerebrum, basal ganglia, and cerebellum to ready the system to produce the phrase in the form of a motor schema leading to activation of muscle groups.

Motor Schema

A rough plan of speech production based upon the abstract representation of the mechanism. General instructions are fed forward in syllable chunks. Instructions are flexible enough to allow for variations.

Response Feedback

Accounts for self-regulation of muscle groups and also reports to schema centers for feedforward predictive control of general instructions.

Articulator Movements and Cavity Changes

Both phoneme and syllable disappear in the quasi-continuous movements involved in producing the phrase. Coarticulatory variations are accounted for by self-regulation within muscle groups.

Air Pressure and Acoustic Output

Air pressure variations within the vocal tract set up audible pressure waves heard as [wibitʃuənsɑkɝ].

External Feedback

Sensations of touch, air pressure, and audition relay information to the speaker about his own speech for self-correction.

**Figure 1.** Model of speech production. A conceptualization of a skilled speaker's production of "We beat you in soccer." Reprinted with permission from G. J. Borden and K. S. Harris, *Speech science primer.* Baltimore, Md.: Williams & Wilkins, © 1980.

close vowel. The constant compensation for change that is evident in speech means either that muscle groups can autonomously regulate themselves, acting together as a "coordinate structure" (Fowler, Rubin, Remez, & Turvey, in press), or that the muscle groups report on their performance to higher centers so that the motor schema can be altered appropriately. This information on the performance of the muscles, called response feedback, is relayed for speech primarily by muscle spindles embedded in the respiratory, laryngeal, and articulatory muscles. Response feedback from muscles is relatively rapid since it precedes any movement that may result. Specialized muscle fibers, the spindles, lying in parallel with the main muscle fibers are thought to report information on muscle length and the rate at which the length is changing (Matthews, 1964). Thus simultaneous reports from a group of cooperating muscles could result in information on the relative contribution of each muscle in the group, allowing adjustments to be made for any imbalance. Constant responsiveness of the spindles is made possible by the simultaneous contraction of the main muscle and the spindle fibers during voluntary movement (Vallbo, 1971). Since the muscle spindles of the tongue lie in a three-directional pattern, they can provide complex three-dimensional information from many muscles simultaneously, yielding information on subtle changes in position and shaping. Cooper (1953) reported spindles to be especially prominent near the midline of the superior longitudinal muscle in a region proximal to the tip, the most flexible part of the tongue.

When the muscle-group cooperatives act to produce movements and changes in vocal cavity shape, the resulting changes in air pressure are made audible and form the sounds of speech. The speaker hears and feels his own speech being produced, and these sensations of touch and audition can be used for self-correction of errors. Auditory and tactile feedback can be classified as external feedback because they arise after the motor patterns for speech have been initiated. For the ballistic, fast-acting gestures of speech, such as the stop consonants, touch and audition occur too late to provide ongoing control of the motor commands. For continuants, however, they may serve to provide finer tuning of the productions to better match the perceptual targets. Tactile feedback is monitored by surface receptors responsive to light pressure changes and by subcutaneous receptors responsive to deep pressure. The anterior superior surface of the tongue is more sensitive to touch than the back or inferior parts of the tongue. In addition, the lips and the alveolar ridge of the palate are sensitive to light touch (Ringel & Ewanowski, 1965; Ruch, 1960) and are important contact sites in speech. For auditory feedback, sound is conducted by air through the ear to the cochlea and also by bone through the skull of the speaker to the cochlea. The bone-conducted sound consists primarily of the low frequencies represented in the speech spectrum, while the air-conducted sound includes both high and low frequencies. However, the air- and bone-conducted sounds are approximately equal in intensity (Békésy, 1949).

## III. EFFECTS OF ALTERED FEEDBACK

During the last 30 years, there have been many studies of normal speakers talking under various conditions of altered feedback. The auditory feedback of the speech signal has been delayed, amplified, attenuated, filtered, and altogether masked. The sense of touch has been diminished by blocking sensation from tongue and palatal receptors via the injection of anesthesia into the appropriate sensory nerves. The normal proprioceptive sense has been altered indirectly by changing the shape of the vocal tract with palatal prostheses or by imposing external constraints upon movement.

In all of these adverse feedback conditions, the result has been that speakers invariably compensate for any perturbation. When they hear their own speech delayed by a fraction of a second, they stall in an apparent attempt to let the auditory feedback catch up (Lee, 1950). When the auditory feedback is normal in timing but increased or decreased in intensity, speakers adjust their vocal intensity to try to match the feedback to their intended vocal effect (Lane & Tranel, 1971; Siegel & Pick, 1974). When the speech is filtered, speakers attempt to restore the missing frequencies (Garber & Moller, 1979). The adjustments made indicate that speakers are using auditory feedback to monitor their speech in situations of distorted feedback. However, whether speakers use audition to monitor themselves to the same degree during normal feedback is not known. Unable to hear themselves when high-level noise is used to mask the air-conducted sound and a bone vibrator is used to mask the bone-conducted sound, speakers continue to be highly intelligible. However, they increase vocal intensity and prolong vowels, trying to normalize the situation.

When oral anesthesia is applied to reduce tactile feedback, speech suffers subtle articulatory distortions, especially prominent in /s/-stop clusters (Borden, Harris, & Oliver, 1973). However, the speech is remarkably intelligible. When auditory and tactile feedback are diminished at the same time, the resulting speech remains intelligible, but the prosodic effects of the auditory masking are added to the articulatory effects of the tactile deprivation. The compensatory strategies that may operate to maintain intelligibility under these circumstances are difficult to measure at present. There is evidence of general muscle reorganization (Borden, Harris, & Catena, 1973; Leanderson & Persson, 1972), evidence of tongue retraction (Scott & Ringel, 1971), and evidence that airflow and air pressure are increased (Hutchinson & Putnam, 1974; Prosek & House, 1975).

It is perhaps easier to observe the compensatory mechanisms at work under conditions in which a bite block is inserted between the teeth, in which there is mechanical resistance applied to an articulatory movement, or, finally, in which the vocal tract is changed by inserting a false palate into the oral cavity. For vowel productions, speakers compensate immediately for a fixed jaw (Lindblom, Lubker & Gay, 1979), apparently by altering tongue movement. When normal jaw movement is resisted mechanically, the lips compensate (Folkins & Abbs, 1975). Alterations of the vocal tract with prostheses change vocal tract coordinates and thus present conditions in which the subject must recalibrate the system; thus, some trial and error is evident in developing compensatory patterns (Borden *et al.*, 1979; Hamlet & Stone, 1976).

Another factor evident from altered feedback studies is related to the compensation detailed above. The fact that the feedback system for speech is highly redundant may account for much of the compensation observed. The system exists on so many levels that it is impossible to eliminate speech feedback altogether. When one channel is blocked, the speaker depends more upon another channel of information. Even in the rare case of adventitious auditory agnosia, the speaker can talk intelligibly despite the fact that all incoming sounds lack significance (Lassen, 1979). Although the sound of his own speech is presumably useless to him, such a patient can rely upon proprioceptive information to monitor his performance.

How does one explain studies of animal movement in which even proprioception is eliminated by severing the sensory roots that provide all afferent information from movement? Animal subjects can continue, without visual feedback, to perform learned limb or jaw movements after complete deafferentation (Goodwin & Luschei, 1974; Polit & Bizzi, 1978; Taub & Berman, 1968). Perhaps the answer lies in the redundancy of the system: the internal feedback loop may be sufficient by itself. The brain knows that its performance is matching its intention, hence no report from the periphery is required.

A final observation on the altered feedback experiments is the variability in the effects upon human subjects. Delayed auditory feedback renders some subjects almost unable to speak, whereas others can fairly successfully ignore the auditory signal. Among subjects given a presumably identical injection of anesthesia to block sensation from the lingual nerve, some will experience a noticeable effect upon their speech, whereas the speech of others will show no perceptual effect (Bordon, Harris & Oliver, 1973). The degree to which this variability may reflect an equally large variability in the methods used to monitor speech under normal conditions can only be inferred.

## IV. FEEDBACK DURING SPEECH ACQUISITION

The critical time for the development of auditory, tactile, and proprioceptive feedback associations is regarded by many to be the period of babbling. Before babbling, the infant makes vegetative, reflexive sounds connected with comfort, discomfort, and hunger. When babbling appears, it is mixed with cooing, but distinguished by its syllable-like repetitions of constricted vocal tract, consonant-like sounds releasing into more open vocal tract, vowel-like sounds. Babbling seems to be preprogrammed in the developmental sequence of motor activities: both deaf babies and hearing babies babble in much the same way. After babbling for some time, the hearing baby presumably begins to attend to it and exert voluntary control over it. The baby hears and feels its own sound productions which become increasingly similar to the language of its caretakers (Cruttenden, 1970; Weir, 1966). This requires the development of associations among the various feedback channels and motor patterns. In contrast, the deaf baby gradually babbles less (Mavilja, 1969) as there is presumably less reward in vocal play: he can hear neither himself nor others.

Normal babies are born with an auditory mechanism that can distinguish between sounds that will, in many cases, be useful for phonemic distinctions later. The work of Eimas, Siqueland, Jusczyk, and Vigorito (1971) and other studies in infant perception have shown that infants only a few weeks old are able to make fine auditory distinctions that correspond to contrasts of manner, place, and voicing in many languages (see Kuhl, 1978, for a review).

As an example of this finding, studies of the distinction between the sounds /r/ and /l/ can be offered. Babies 2 and 3 months old were tested on speechlike stimuli varying in frequency change of the third formant, the most prominent acoustic cue used to distinguish /r/ from /l/. The sounds were presented contingent on the baby's sucking a pacifier. (With repetition of what the baby perceives as the same stimulus, sucking rate gradually decreases; but if a sound is introduced that the baby perceives as novel, sucking rate increases. Thus, a change in sucking rate indicates discrimination between habituated and novel sounds.) Infant discriminations were tested between stimuli that adults hear as /r/ and /l/, between stimuli that differ acoustically but are both heard as /r/ by adults, between stimuli also differing in $F_3$ change but heard by adults as /l/, and between identical stimuli. Infants gave evidence of a reliable increase in sucking rate when the stimuli paired were those heard by adults to be /r/ and /l/ as opposed to the other pairs (Eimas, 1975). Thus, babies seem to hear /r/–/l/ differences, among many others, and in their early babbling /r/- and /l/-like sounds are frequent (Clark & Clark, 1977); but during the second, more volitional stage of babbling, /r/ and /l/ drop out (Jakobson, 1968). The increasing use of voluntary control in babbling, although important in developing motor control, thus produces some interesting changes in the vocal repertoire.

As babbling interweaves with the beginnings of first words, another aspect of production and perception is added. Not only has the child assumed some voluntary and increasingly precise control over sound production, but these sounds have started to take on meanings. In the production of meaningful utterances /r/ and /l/ appear late (Templin, 1957; Jakobson, 1968). Furthermore, when embedded in a linguistic task, perception of /r/ and /l/ is also late (Shvachkin, 1973). It seems that acoustic features important to language are detected soon after birth, but speech perception in context takes time to develop (Edwards, 1974; Zlatin & Koenigsknecht, 1975). It may also be that feedback sufficient to gain control over relatively simple oral sound productions develops early, but the feedback mechanisms and motor control necessary to produce complex speech patterns in which semantic processes are involved develop slowly. Evidence of this slowly developing control comes from recordings of children practicing their newly discovered language by themselves. The repetitions and variations of phrases resemble phonetic practice as much as they do vocabulary practice (Weir, 1966, pp. 166–167). In the following sample, Weir's child was practicing variations in initial consonant:

(1) fumbelina (2×)
(2) tumbelina
(3) lumbelina
(4) Thumbelina (2×)

In developing control of speech, children seem to use adult models as their perceptual targets, but their perception of the targets is apt to be rather undifferentiated and their ability to reproduce the targets imprecise and variable. In general, their perceptual ability seems to develop faster than their production ability. In identification and reproduction tasks on sets of synthetic consonant–vowel–consonant (CVC) syllables in continua from *light* to *white*, from *light* to *write*, and from *white* to *write*, 4-year-old children could identify the syllables better than they could repeat them (Menyuk & Anderson, 1969).

Strange and Broen (in press) compared the perception and production of /r/–/l/ along with other semivowels in 3-year-old children and found that they were good at differentiating such contrasts as *rake* and *lake* and that they perceived the phoneme boundary in a synthetic series much as college students do. Their most interesting finding was that children who have mastered the /r/–/l/ contrast in their speech, or who showed few distortions, were good at identifying clear examples of the contrast—whether it was delivered by live voice or by tape-recorded voice or was synthetically produced. However, among children with many distortions and substitutions in their speech, half made very few identification errors but half had some difficulty perceiving the contrasts. Thus, there does seem to be some relationship, between perception and production, but even in the worst cases, perception was better than production.

The relationship between a child's perception and production is complex. It changes in mysterious ways as the child develops. [A. Gerber (personal communication) of Temple University illustrates this with a language sample from a child named Eric. At 14 months, Eric called a dog "fa-fa." His mother would say "See the dog? Ruff-ruff. That's a ruffy dog." Eric would say "fa-fa." By 19 months, he was saying "goggy" when he pointed to a dog. His grandmother said to him "You used to call him a 'fa-fa', whereupon Eric said "Ruffy gog."]

Although in advance of production, perception does seem to be developmental. Chaney and Menyuk (1975) reported that 4-year-olds who produced [w] for /r/ and /l/ and 6-year-olds who produced /r/ errors were not as accurate as controls with good articulation in pointing to pictures of *light, write,* or *white* when adults produced them; furthermore, the 4-year-olds could only identify between 33 and 46% of their own tape-recorded productions.

This question of how children perceive themselves during their parallel refinement of perception and production is of interest to those theorizing on feedback and motor control (see McReynolds, 1978, for a review). In the aforementioned Chaney and Menyuk study, children heard themselves on a tape recording and failed to point to the picture that they had previously named. Also, Locke and Kutz (1975) found that 5-year-old children who say "wing" for *ring* can perceive the /r/ correctly in adult speech, but when they hear their own tape-recorded misarticulations of *ring,* they are more likely to point to a picture of a wing than a ring.

On the other hand, some people report that children may be hearing some difference in their own misarticulations that adults fail to hear. Support for this is largely anecdotal.

Child: "She's wearing a wing."
Adult: "A wing?"
Child: "No, not a bird wing, a wedding wing."

Kornfeld (1971) found a spectrographic difference between the [w] in a child's production of [gwæs] for *glass* and the [w] in the same child's production of [gwæs] for *grass*. This finding suggests that there may be some differences in production that are not phonemically significant to adults.

The apparent discrepancy between the notion that children are perceiving a difference that adults fail to perceive and the view that children usually perceive much as adults do, but lack production proficiency, may be partly resolved if we realize that the anecdotal support for the first notion comes from children's making discriminations of their own utterances while speaking; but support for the second idea comes from children's discrimination of others speaking or of their own tape-recorded speech. During live speech, children can feel themselves as well as hear themselves. Thus, they may sense small differences between *wing* for bird and wedding *wing* that neither they nor adults perceive as different by audition alone. One could infer either that the children were aware of a difference in intention or that there was some phonemically insignificant difference in muscle activity that they could sense. Spectrographic differences support the second alternative. The logic of this explanation might lead us to the conclusion that in these cases, children are using proprioception more than audition to monitor themselves, or, more conservatively, that muscle sense is needed in addition to the hearing sense for some children to distinguish their own /w/–/wʳ/ distinctions.

Development of feedback mechanisms has been difficult to study. Studies of delayed auditory feedback (DAF) on the speech of children have shown older children (7 to 9 years) to be more affected by the delay than are younger children (4 to 6 years) (Chase, Sutton, First, & Zubin, 1961), but later studies found younger children to be more affected than are older children or adults (MacKay, 1968). Using DAF to test the development of auditory feedback may not be the best method because the technique forces attention to audition and creates an artificial discrepancy between muscle information and the auditory information returned to the speaker (Borden, Dorman, Freeman, & Raphael, 1976). Studies of children speaking in noise, and while hearing their own voices amplified, show that the children alter their vocal intensity appropriately but compensate to a different degree than do adults. The children reduced vocal intensity less under amplified feedback than adults did, but did not differ significantly in the degree to which they increased vocal intensity when speaking in noise (Siegel, Pick, Olsen & Sawin, 1976). By 4 years of age, children compensate much as adults do for any artificial interference in feedback. For example, nerve-block anesthesia of the tongue resulted in the same proportion of affected articulation in 4-year-olds as was found in adults (Borden, 1976). The redundancy of the feedback systems and the inaccessibility of the proprioceptive system to intervention prevent a direct comparison of the different feedback systems. A predictive feedforward system with CNS control may be fairly well developed by 4 years. Also, **muscle sense** may take over when the auditory or tactile senses are diminished. Fortunately, these problems for the researcher are assets for the speaker.

We are left with the realization that we know very little about the development and use of feedback control mechanisms in children who are learning to speak. By the time children are old enough to be tested, they are fairly skilled speakers, despite the fact that some speech refinements are still being acquired. We know, however, that children are much more competent than are adults in learning new languages. There is an apparent plasticity of the neurological networks that persists until approximately the age of puberty (Lenneberg, 1967; Penfield & Roberts, 1959). It behooves us to test children on unfamiliar or more recently learned items as well as on well-known utterances—and to compare young children, older children, and adults on their relative dependence upon feedback when producing speech gestures varying in familiarity.

## V. CONTROL OF ESTABLISHED SPEECH

There have been some studies of altered feedback in adults in which the effects were studied on the speaker's native language, a second language, and an entirely novel language. One such study that was well controlled and designed (MacKay, 1970) tested 21 English–German bilinguals, for some of whom German was the native language and for others of whom English was the native language. Subjects read 15-syllable sentences in both languages and in a completely unfamiliar language, Congolese. In the experimental condition, the subjects read sentences under DAF. Native speakers of German made more stuttering-like responses when speaking English than when speaking German, and native speakers of English made more stuttering-like responses when speaking German. Both groups of speakers had more difficulty with the unfamiliar language, Congolese. Thus, DAF was more of an interference to the less familiar language and most disruptive of the novel Congolese sentences.

In an attempt to control the variable of attention, MacKay had subjects read in German and in English while listening to their own voices speaking the opposite language. They spoke both native and second languages more slowly with this distraction, but it interfered more with their native language. MacKay suggests that this may have been caused by increased concentration on the articulation of the less familiar language. A slight variant of this explanation might be that the native language could be put to some degree on automatic control so that subjects could do two things at once: speak and listen to an irrelevant voice; whereas for the second language, they had to maintain more control and could not, at the same time, attend to an unrelated voice.

Why, then, would DAF interfere more with the less familiar language? Would not MacKay's suggestion hold here as well? I am persuaded that the important variables are the intensity and nature of the auditory information. In the DAF condition, the distorted signal was delivered at 95-dB sound pressure level (SPL) and the speaker recognized the signal as his own ongoing speech, delayed. The irrelevant voice condition was not amplified to such an extent and, even if recognized by the speaker, was not associated with the ongoing speech. Thus, when the speaker is most skilled, as in speaking his or her native language, less control is needed and there is less interference from a distortion of feedback; but, for the same reason, the speaker is free to attend to something unrelated to the speech.

One barrier to an interpretation of DAF studies has been that DAF is described as interference with auditory feedback. The crux of the DAF effect is not just that the auditory feedback is delayed, but that the auditory signal is amplified beyond ignoring and presents competing information with the proprioceptive feedback within the speaker. Delayed auditory feedback presents the subject with an auditory–proprioceptive mismatch.

We get a glimpse of how speech may be controlled by skilled speakers and how perception and production may interact when we study speakers who are learning a new language. In a study currently underway at Haskins Laboratories, we are finding that skilled speakers produce unfamiliar gestures in accordance with their perceptions of them. One speaker will perceive the sounds as novel and use trial-and-error strategies to match the unfamiliar target; another speaker, also perceiving the sounds as novel, will incorporate any new sound into his own production system and make unfamiliar items familiar in terms of motor patterns. The difference in the way subjects link the perceptual target with the production target may determine the difference between utterances by speakers imitating syllables under various conditions of feedback deprivation. The speaker attempting to match a novel perceptual target, such as /χi/, shows much more variability on unfamiliar gestures than on familiar gestures; but the speaker adhering to familiar perceptual targets, such as perceiving /χi/ as a variant of /ʃi/ or /hi/, displays variability under various feedback conditions—but no more variability on unfamiliar than on familiar gestures. This investigation also shows that, even for skilled speakers producing familiar speech sounds, feedback plays a part in the fine-tuning of the match between the actual production and the perceptual target. For example, the second formant of vowels was found to be higher when subjects are able to hear themselves, whether the vowel is the familiar /i/ or the less familiar /y/ (Borden *et al.*, 1979).

Perception thus has a strong influence on production. It is also true that production, or language experience, affects perception of speech sounds. When listeners are presented with a series of synthetic sounds that form an acoustic continuum between two phonemes, each member of the series differing from the next by an equal acoustic change, there is a strong tendency for the sounds to be heard categorically. That is, some members of the series are heard as one phoneme, some as the other phoneme, and discrimination between members of the series is high at the phoneme boundary but low among the sounds labeled as a particular phoneme (Liberman, Harris, Hoffman, & Griffith, 1957).

It has been demonstrated that linguistic experience, as well as psychoacoustic ability, affects adult perception of phonemic boundaries along such acoustic continua. The positive influence of production upon perception has been evident in an increased ability to discriminate acoustic dimensions related to linguistic experience and a corresponding decreased ability if the distinction is not related to linguistic experience (see Strange & Jenkins, 1978, for a review). For example, one study (Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975) compared speakers of Japanese with speakers of American English on their perception of the liquids /r/ and /l/. These sounds are contrastive in English, as in *rake* and *lake*, but not contrastive in Japanese. Thus, American English listeners divided a series of synthetic syllables varying in $F_3$ starting frequency into two categories: those that sounded more like /ra/ and those that sounded more like /la/. Listeners from Japan tended to hear all the stimuli as /ra/. To test discrimination, listeners were presented with a series of three items from the set: two of them identical and one different. Subjects indicated whether the different syllable was heard first, second, or third. American listeners could tell the difference 80% of the time if the compared stimuli were near the /r/–/l/ boundary, but their discrimination scores fell to between 40 and 60% if the stimuli were those that they identified as a single phoneme. The Japanese listeners, in general, showed no sharp discrimination peak. In discrimination of nonspeech stimuli consisting of $F_3$ without the other formants, Japanese and American groups both discriminated between 66 and 89% correct. It appears that subjects report auditory differences among nonspeech sounds equally well; but, when listening to speechlike sounds, they report only the differences important in their own speech.

The Miyawaki *et al.* (1975) study used, as Japanese subjects, students living in Japan. These subjects had studied English in school but had little practical experience in speaking English. Another study of a Japanese group and an American group of speakers living in Japan pointed to a more complex relationship between perception and production (Goto, 1971). Both groups were recorded producing words containing /r/ and /l/, such as *collect*, *correct*, *play*, and *pray*. Then each subject was asked to identify either /r/ or /l/ in the words as recorded by themselves and by the others. American speakers were good at identification of /r/ and /l/ in their own speech and in the speech of Japanese speakers. Some of the Japanese subjects produced the contrast to the satisfaction of American listeners and some did not; but the Japanese listeners were poor at identification, whether they could produce the contrast or not. Goto concluded that the Japanese speakers good at production but poor in identification must be using kinesthetic feedback. Some of the Japanese subjects were tested further on a combined identification–discrimination task. They responded "r–l different" or "l–l same" to pairs of words, *pray–play* or *collect–collect*, as recorded by American speakers. Inspection of the data presented indicates that discrimination was more difficult than identification for the Japanese. Goto tested himself on the identification task after 1 month of studying English conversation, after 4 months, and after 10 months. He reported little improvement in his perception, but no mention was made of whether his speech improved during that time. The Goto study has perplexed many investigators because it seems to indicate that correct production of a phonemic contrast can be developed in the absence of any auditory perception of the difference. It may be, however, that despite poor discrimination, some degree of auditory identification is necessary initially to associate the new production with some sound of the speaker's first language. Both production and perception of a novel phoneme in the second language may be related to a variation of an item in the first language. It follows that whatever feedback and feedforward mechanisms operate to control the sound patterns of the first language are quickly adopted for the second. It would be interesting to track the relative importance of feedback information as skilled speakers produce their native speech patterns and as they learn new patterns.

## VI. CONCLUSION

We have reviewed evidence for at least three levels of information flow that can be used to direct the production of speech: internal feedback, response feedback, and external feedback. Internal feedback is made possible by circuits among the cerebrum, midbrain, and cerebellum of the brain. Response feedback from much of the speech musculature is relayed from muscle spindles, and, added to any information from tendon and joint receptors, is thought to form the complex sensation of movement and position. External feedback through the auditory and tactile systems yields knowledge of results to the speaker, knowledge that can be used for fine-tuning of the speech signal or for the correction of errors.

Experimental interference with external feedback has demonstrated the remarkable compensatory abilities of speakers in general, but has also pointed to large variation in effects among individual speakers.

As children acquire speech, they seem to build upon their innate acuity in detection of speech sounds; but perception of distinctions embedded in meaningful stimuli evidently develops with age along with, and somewhat in advance of, their production abilities. Self-perception must play an important role in forming associations between speech perception and speech production in speakers acquiring new speech patterns. Skilled adult speakers perceive speech in accordance with their linguistic experience and are generally more inhibited in producing new patterns than are children.

Learning a motor skill seems to require some knowledge of results (Adams, 1971), and children who engage in variable practice can adapt quickly to learn new tasks (Kelso & Norman, 1978). Children acquiring their first language engage in variable practice and seem to depend upon their well-developed auditory discrimination to refine their speech and to give them knowledge of their progress. To what degree the other feedback mechanisms are used can only be imagined at present. During the critical period of speech learning (Lenneberg, 1967; Marler, 1975), children display a tendency toward experimentation and the active use of feedback. It does seem, however, that perceptual representations and articulation programs in young children are both unstable; that perception stabilizes ahead of motor control; and, thus, that older children are better able to benefit from feedback as they gain motor proficiency (Newell & Kennedy, 1978). Many of the references cited in the foregoing refer to the development of motor learning in areas other than speech. We must study this development more closely in speech: Do we indeed depend less upon feedback as we become more skilled as speakers? Preliminary studies indicate that this may be the case.

The strategies used in learning a new phonetic system may depend upon whether the speaker is still within the critical period for language learning or well beyond it. There are indications that children learning a second language keep the feedback channels open. A study of Puerto Rican children learning English (Williams, 1974) showed perception to improve as experience with the new language increased, but the younger children changed more rapidly toward English perceptual boundaries than did the older children. The increased sensitivity to the contrasts important in the new language was found to interfere temporarily with the native language.

Adults learning a new language, however, seem to base the new language on their native language. They use feedback, but lacking established links between feedback and the new production programs, their new speech gestures tend to be modifications of their old system, complete with whatever degree of automaticity is involved in that system. Skilled speakers are good at perceiving distinctions important in their native language (Miyawaki *et al.,* 1975), but poor at perceiving distinctions important in a less familiar language, even if they produce the distinctions passably well (Goto, 1971). Only to the degree that adults can ignore their previously learned sound system and can become childlike in their freedom to experiment and in sensitivity to their own productions will they enjoy success in achieving the suprasegmental and segmental nuances of a new language.

## Acknowledgment

The writing of this chapter was partially supported by NIH grant NS13870.

## References

Adams, J. A. A closed-loop theory of motor learning. *Journal of Motor Behavior,* 1971, 3, 111–150.

Békésy, Georg von. The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America,* 1949, 21, 217–232.

Borden, G. J. An interpretation of research on feedback interruption in speech. *Brain and Language,* 1979, 7, 307–319.

Borden, G. J. The effect of mandibular nerve block upon the speech of four-year-old boys. *Language and Speech,* 1976, 19, 173–178.

Borden, G. J., Dorman, M. F., Freeman, F. J., & Raphael, L. J. Electromyographic changes with delayed auditory feedback of speech. *Journal of Phonetics,* 1976, 4, 1–8.

Borden, G. J., & Harris, K. S. *Speech science primer.* Baltimore: Williams & Wilkins, 1980. Pp. 155–156.

Borden, G. J., Harris, K. S., & Catena, L. Oral feedback. II. An electromyographic study of speech under nerve-block anesthesia. *Journal of Phonetics,* 1973, 1, 297–308.

Borden, G. J., Harris, K. S., & Oliver, W. Oral feedback. I. Variability of the effect of nerve-block anesthesia upon speech. *Journal of Phonetics,* 1973, 1, 289–295.

Borden, G. J., Harris, K. S., Yoshioka, H., & Fitch, H. *Feedback and feedforward mechanisms used by speakers producing familiar and novel speech patterns*. Paper presented at the 9th International Congress of Phonetic Sciences, Copenhagen. 1979.

Chaney, C. F., & Menyuk, P. *Production and identification of /w,r,l,j/ in normal and articulation impaired children*. Paper presented at the Annual Convention of the American Speech & Hearing Association, Washington, D.C., 1975.

Chase, R. A., Sutton, S., First, D., & Zubin, J. A developmental study of changes in behavior under delayed auditory feedback. *Journal of Genetic Psychology*, 1961, **99**, 101–112.

Clark, H. H., & Clark, E. V. *Psychology and language*. New York: Harcourt Brace Jovanovich, 1977.

Cooper, S. Muscle spindles in the intrinsic muscles of the human tongue. *Journal of Physiology*, 1953, **122**, 193–202.

Cruttenden, A. A phonetic study of babbling. *British Journal of Disorders of Communication*, 1970, **5**, 110–117.

Edwards, M. L. Perception and production in child phonology: The testing of four hypotheses. *Journal of Child Language*, 1974, **1**, 205–219.

Eimas, P. D. Auditory and phonetic coding of the cues for speech: Discrimination of the r-l distinction by young infants. *Perception & Psychophysics*, 1975, **18**, 341–347.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. Speech perception in infants. *Science*, 1971, **171**, 303–306.

Evarts, E. V. Feedback and corollary discharge: A merging of the concepts. *Central Control of Movement: Neurosciences Research Program Bulletin*, 1971, **9**, 86–112.

Folkins, J. W., & Abbs, J. H. Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *Journal of Speech & Hearing Research*, 1975, **18**, 207–220.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production*. New York: Academic Press, in press.

Fry, D. B. The control of speech and voice. In H. Kalmus (Ed.), *Regulation and control in living systems*. New York: Wiley, 1966. Pp. 137–154.

Garber, S. R., & Moller, K. T. The effects of feedback filtering on nasalization in normal and hypernasal speakers. *Journal of Speech and Hearing Research*, 1979, **22**, 321–333.

Goodwin, G. M., & Luschei, E. S. Effects of destroying the spindle afferents from the jaw muscles upon mastication in monkeys. *Journal of Neurophysiology*, 1974, **37**, 967–981.

Goto, H. Auditory perception by normal Japanese adults of the sounds "L" and "R." *Neuropsychologia*, 1971, **9**, 317–323.

Hamlet, S. L., & Stone, M. Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses. *Journal of Phonetics*, 1976, **4**, 199–218.

Hutchinson, J. M., & Putnam, A. H. B. Aerodynamic aspect of sensory deprived speech. *Journal of the Acoustical Society of America*, 1974, **56**, 1612–1617.

Jakobson, R. *Child language, aphasia, and phonological universals*. The Hague: Mouton, 1968.

Kelso, J. A. S., & Norman, P. E. Motor schema formation in children. *Developmental Psychology*, 1978, **14**, 153–156.

Kimura, D. Acquisition of a motor skill after left-hemisphere damage. *Brain*, 1977, **100**, 527–542.

Kornfeld, J. Theoretical issues in child phonology. In *Proceedings of the seventh annual meeting of the Chicago Linguistic Society* (CLS7). University of Chicago, 1971, 454–468.

Kuhl, P. K. Predispositions for the perception of speech-sound categories: A species-specific phenomenon? In F. D. Minifie & L. L. Lloyd (Eds.), *Communicative and cognitive abilities: Early behavioral assessment.* (NICHD Series) Baltimore: University Park Press, 1978. Pp. 229–255.

Lane, H. L., & Tranel, B. The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research,* 1971, **14,** 677–709.

Lassen, N. A. *The physiology and pathophysiology of language functions as illustrated by measurements of the regional blood flow in the cortex of the brain.* Special lecture presented at the 9th International Congress of Phonetic Sciences, Copenhagen, 1979.

Leanderson, R. & Persson, A. The effect of trigeminal nerve block on the articulatory EMG activity of facial muscles. *Acta Otolaryngologica,* 1972, **74,** 271–278.

Lee, B. S. Effects of delayed speech feedback. *Journal of the Acoustical Society of America,* 1950, **22,** 824–826.

Lenneberg, E. H. *Biological foundations of language.* New York: Wiley, 1967.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C., The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology,* 1957, **54,** 358–368.

Lindblom, B., Lubker, J., and Gay, T. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics,* 1979, **7,** 147–161.

Locke, J. L., & Kutz, K. J. Memory for speech and speech for memory. *Journal of Speech and Hearing Research,* 1975, **18,** 176–191.

MacKay, D. G. Metamorphosis of a critical interval: Age-linked changes in the delay in auditory feedback that produces maximal disruption of speech. *Journal of the Acoustical Society of America,* 1968, **43,** 811–821.

MacKay, D. G. How does language familiarity influence stuttering under delayed auditory feedback? *Perceptual and Motor Skills,* 1970, **30,** 655–669.

MacNeilage, P. F. Motor control of serial ordering of speech. *Psychological Review,* 1970, **77,** 182–196.

Marler, P. On the origin of speech from animal sounds. In J. F. Kavanagh & J. E. Cutting (Eds.), *The role of speech in language.* Cambridge, Mass.: MIT Press, 1975, pp. 11–37.

Matthews, P. B. C. Muscle spindles and their motor control. *Physiological Review,* 1964, **44,** 219–288.

Mavilja, M. P. *Spontaneous vocalization and babbling in hearing impaired infants.* Unpublished doctoral dissertation, Columbia University, 1969.

McReynolds, L. V. Behavioral and linguistic considerations in children's speech production. In J. F. Kavanagh & W. Strange (Eds.), *Speech and language in the laboratory, school, and clinic.* Cambridge, Mass.: MIT Press, 1978, pp. 127–164.

Menyuk, P., & Anderson, S. Children's identification and reproduction of /w/, /r/, and /l/. *Journal of Speech and Hearing Research,* 1969, **12,** 39–52.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics,* 1975, **18,** 331–340.

Newell, K. M., & Kennedy, J. A. Knowledge of results and children's motor learning. *Developmental Psychology,* 1978, **14,** 531–536.

Penfield, W., & Roberts, L. *Speech and brain-mechanisms.* Princeton, N.J.: Princeton University Press, 1959.

Polit, A., & Bizzi, E. Processes controlling arm movements in monkeys. *Science,* 1978, **201,** 1235–1237.

Prosek, R. A., & House, A. S. Intraoral air pressure as a feedback cue in consonant production. *Journal of Speech and Hearing Research*, 1975, **18**, 133–147.

Ringel, R. L., & Ewanowski, S. J. Oral perception: Two-point discrimination. *Journal of Speech and Hearing Research*, 1965, **8**, 389–398.

Ruch, T. C., & Fulton, J. F. *Medical physiology and biophysics*. Philadelphia: W. B. Saunders, 1960.

Scott, C. M., & Ringel, R. L. Articulation without oral sensory control. *Journal of Speech and Hearing Research*, 1971, **14**, 804–818.

Sherrington, C. S. *The integrative action of the nervous system*. London: Constable, 1906.

Shvachkin, N. K. The development of phonemic speech perception in early childhood. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development*. New York: Holt-Rinehart, 1973. Pp. 91–127.

Siegel, G. M., & Pick, H. L., Jr. Auditory feedback in the regulation of the voice. *Journal of the Acoustical Society of America*, 1974, **56**, 1618–1624.

Siegel, G. M., Pick, H. L., Jr., Olsen, M. G., & Sawin, L. Auditory feedback in the regulation of vocal intensity of preschool children. *Developmental Psychology*, 1976, **12**, 255–261.

Stevens, K. N., & Perkell, J. S. Speech physiology and phonetic features. In M. Sawashima & F. S. Cooper (Eds.), *Dynamic aspects of speech production*. Tokyo: University of Tokyo Press, 1977. Pp. 323–341.

Strange, W., & Broen, P. A. Perception and production of approximant consonants by three-year-olds: A first study. In G. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Perception, production and deviation*. Proceedings of a NICHHD conference, Bethesda, Md., May, 1978, in press.

Strange, W., & Jenkins, J. J. Role of linguistic experience in the perception of speech. In R. D. Walk & H. L. Pick, Jr. (Eds.), *Perception and experience*. New York: Plenum, 1978. Pp. 125–169.

Taub, E., & Berman, A. J. Movement and learning in the absence of sensory feedback. In S. J. Freedman (Ed.), *The neuropsychology of spatially oriented behavior*. Homewood, Ill.: Dorsey Press, 1968. Pp. 173–192.

Templin, M. *Certain language skills of children*. Minneapolis: University of Minnesota Press, 1957.

Vallbo, Å. B. Muscle spindle response at the onset of isometric voluntary contractions in man. *Journal of Physiology*, 1971, **318**, 405–431.

Weir, R. H. Some questions on the child's learning of phonology. In F. Smith & G. A. Miller (Eds.), *The genesis of language*. Cambridge, Mass.: MIT Press, 1966. Pp. 153–168.

Williams, L. *Speech perception and production as a function of exposure to a second language*. Unpublished doctoral dissertation, Harvard University, 1974.

Zlatin, M. A., & Koenigsknecht, R. A. Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research*, 1975, **18**, 541–553.