ACQUISITION OF SPEECH PRODUCTION: THE ACHIEVEMENT OF SEGMENTAL INDEPENDENCE

PETER F. MACNEILAGE & BARBARA L. DAVIS
The University of Texas at Austin
Austin Tx. 78712
U.S.A.

ABSTRACT. As a primate communicative event, the repetitive, rhythmic, open-close alternation of the mandible, accompanied by phonation, is observable in three forms; some variants of the lipsmack, which is widespread in other primates, the initial babbling of human infants, and the production of the syllables of adult speech. In the first two of these, successive cycles tend to be uniform but in the third there is a highly variegated pattern in successive cycles. Adult segmental serial ordering errors, (e.g. spoonerisms) the effects of which are strongly constrained in terms of syllable structure, suggest that variegation is achieved by placement of independently controlled "Content" elements in syllable "Frames". This paper considers implications of the view that development of infant communicative vocalizations from initial reduplicated babbling to variegated babbling and then to speech, primarily involves a gradual functional differentiation of segmental and subsegmental content elements from a phylogenetically prior basis consisting of 'pure' syllable frames; that is, mandibular oscillations without internal articulatory modulation.

1. Background

Serial ordering errors in normal adult speech that involve single speech segments moving around in an otherwise correct utterance tell us beyond doubt that the individual segment is an independent unit in the control of adult speech. However, there is no evidence that when infants first sound as if they are producing speech segments with speech-like durations, they already have segmental units under independent control. The question addressed in this paper is simple; how do infants achieve this segmental independence?

We wish to address this question in an evolutionary perspective. The independent control of the ordering of segments in a vocal utterance is a new development in primate evolution, a development specific to hominids. However vocalization itself is far from new. Consequently there has been a phylogenetically recent development of superimposition of segmental permutations on vocal output. A clue to the way that this may have been achieved in evolution comes from a well known set of constraints on segmental speech errors known as syllable-position constraints. The constraints are best observed in spoonerisms, but also apply in similar fashion to other serial ordering errors. In spoonerisms, it is almost always true that initial consonants exchange with other initial consonants, vowels exchange with vowels, and final consonants exchange with final consonants. The existence of these syllable position constraints on speech errors gives rise to the "Frame/Content" metaphor for speech motor control; speech production includes a
stage of insertion of segmental "Content" units into syllable structure "Frames" (MacNeilage, Studdert-Kennedy and Lindblom, 1985; MacNeilage, 1987).

The syllable can be described as the interface between the segmental and the suprasegmental levels of speech production. At the suprasegmental or prosodic level, the syllable serves as the unit of rhythmic organization, and the point of reference for the intonation contour and semantic markings. At the segmental level, as we have seen, it serves as a structured repository controlling the serial position of individual segments. If the capacity for vocalization and thus, basically, the capacity for suprasegmental control is not new in hominids, but the segmental capacity is new, how did this change come about? How did the syllable evolve as the interface between the two levels?

Intensive study of the neural mechanisms of monkey vocalization (summarized by Jurgens, 1987) has shown that vocal output is controlled by a system that is medially placed in the neuraxis involving the cerebral cortex and subcortical centres in the midbrain and lower brain stem. At the 'head' of this system are two regions of medial cortex; 1. the anterior cingulate gyrus (ACG), and 2. a neocortical outgrowth of the ACG, the Supplementary Motor Area (SMA). Lesion studies have shown that in monkeys there is no major role of lateral frontal cortex - either lateral premotor cortex, analogous to Broca's Area, or primary motor cortex of the precentral gyrus - in vocalization control. (Jurgens, Kirzinger and von Cramon, 1982).

Important evidence regarding the evolutionary progression from monkey to modern hominids comes from the study of the SMA in humans. A number of different investigators, most notably Penfield and his collaborators (e.g. Penfield and Welch, 1951) have reported evoking rhythmic syllable-like vocalization from electrical stimulation of the SMA. This has not been reported from any other cortical site, to our knowledge. The possibility that these responses reflect the natural function of this area is strengthened by 8 different reports of a total of 17 patients with irritative lesions of the SMA who also involuntarily produced rhythmic syllable-like vocalization (for a summary see Jonas, 1981). These findings lead us to conclude that the SMA plays a leading role in the generation of syllable frames for speech.

A possible precursor to this function in other primates is the well-known lipsmack. This communicative gesture is extremely widespread in higher primates (Redican, 1975). It is a rhythmic series of open-close alternations of the mandible, with the noise that gives it its name being produced either by the release of lip closure or tongue contact with the roof of the mouth. When accompanied by phonation, it is known as the girney (Green, 1975). It seems appropriate that the SMA would serve in an interfacing role between older and newer vocal capacities as it is at once part of the older vocalization control system but unique in that system in being neocortex just as the perisylvian cortex most responsible for language is neocortex. To complete this picture from the motor control standpoint, we would propose that a major event in the evolution of the capacity required for the context-dependent production of segmental content elements in speech is the development of lateral frontal cortex, particularly Broca's Area. (For an expanded discussion of these and other points regarding the neural control of speech production see Lindblom, MacNeilage and Studdert-Kennedy, forthcoming).

As to the relevance of this background for speech acquisition, we suggest, in brief, that an infant recapitulates the phylogenetic steps, beginning without frames or context, developing frames at the onset of babbling, and, much later, following a developmental progression, achieving independent segmental content. The intent of this paper is to see how much of speech acquisition can be understood in terms of a distinction between the role of frames and content elements in the developmental process.