970

In: Bell-Berti, F., & Raphael, L. J. (1995). Producing Speech: Contemporary Issues. For Katherine Safford Harris. AIP Press: New York.

6

Anticipatory and Carryover Effects: Implications for Models of Speech Production

Fredericka Bell-Berti
St. John's University
Haskins Laboratories

Rena A. Krakow
Temple University
Haskins Laboratories

Carole E. Gelfer
William Paterson College

Suzanne E. Boyce
Boston University
M. I. T.

Katherine Harris has been our mentor since the days we each met her. We have decided that one way of thanking Kathy with this volume is by providing a chapter aimed at the new generation of Speech Science students that can serve as a tutorial on coproduction, a model that Kathy helped to formulate. To the extent that this and the rest of our work contributes to increased understanding of speech behavior, Kathy must be credited—for having questioned and deliberated and forced us to think and explain more clearly. We hope that this chapter reflects our mentor's fine hand.

INTRODUCTION

The term "coarticulation" is used to describe the fact that speech is produced as a sequence of sounds in a smooth flow of articulatory movements—that is, there is "blurring of the edges" of segmental articulations as the vocal tract moves from one articulatory configuration to the next. Since studies of speech production have shown the vocal tract to be nearly always on the move from one segment to another, with occasional periods in which some, but not all, articulators can be seen to maintain static positions, it seems obvious that articulatory and acoustical patterns for any one segment must reflect characteristics of (at least) its adjacent segments.

A major goal of speech research has been to discover the spatial-temporal domain of coarticulatory influences, in the belief that they reflect the domain over which the motor system organizes production of a sequence of segmental articulations.[1] By this logic, the observed extent of coarticulatory influence provides us with an estimate of the size of the organizational units of speech production, units whose nature has been much discussed for many years. One problem in establishing the temporal and spatial domains of coarticulatory influence has been that the literature has appeared to provide evidence simultaneously for at least two types of conflicting models of articulatory organization (i.e., "feature-spread" and "coproduction"), in part, we believe, because of a failure to control for intrinsic segmental movements (e.g., Gelfer, Bell-Berti, & Harris, 1989). A second problem is that the studies have assumed that anticipatory and carryover coarticulation are produced via fundamentally different mechanisms, with anticipatory coarticulation envisaged as cognitively controlled, intentional, and large-scale, while carryover coarticulation has been viewed as the small-scale effect of mechanical and inertial forces acting on the articulators (e.g., Lindblom, 1963; MacNeilage, 1970), even in the face of evidence that both carryover and anticipatory coarticulation can occur over similar time spans (Daniloff & Hammarberg, 1973; Fowler, 1984). A third problem has been the failure to acknowledge that coarticulation results from both intersegmental interaction and higher-level prosodic organization. In a 1977 review paper on coarticulation research and models, Kent and Minifie concluded that none of the models of the time were adequate, because none of the theories that were segmentally based could handle the range of reported segmental

---

[1] As developing technologies made it possible to study the production of natural speech, the pervasive nature of coarticulation was viewed as problematic because it contradicted researchers' intuitions that the speech stream is composed sequences of (relatively) invariant segments; it had been, after all, in the domain of speech production that researchers had finally expected to find the invariant units of speech that had eluded them in the acoustical domain. Continuing to maintain the notion that speech is segmentally structured, in the face of the apparently extended range of coarticulatory effects, required an account of speech production that takes a string of discrete segments as its input and outputs a stream of overlapping and asynchronous gestures.

effects, that few attempted to include prosodic effects, and those that did lacked the specificity necessary to allow assessment of their predictions. Although many years have elapsed since the review by Kent and Minifie, their paper remains important for the insights it offers and the gaps it highlights, as the theories continue to require modification and improvement.

As we consider coarticulation research and theory in the mid-1990s, it seems clear that important progress has been made in the intervening years, most notably perhaps in our understanding of the nature of intersegmental organization (e.g., Bell-Berti & Harris, 1981; Browman & Goldstein, 1991; Fowler, 1980; Fowler & Saltzman, 1993). Research has increasingly shown the importance of prosody and syllable organization in predicting patterns of articulatory overlap (e.g., Beckman, Edwards, & Fletcher, 1992; Browman & Goldstein, 1991; de Jong, 1991; Fujimura, 1990; Krakow, 1993; Nittrouer, Munhall, Kelso, Tuller, & Harris, 1988; Turk, 1994; Vaissière, 1988). However, we still lack an adequately detailed model of the combined effects of segmental, syllabic, and prosodic organization.

This chapter reviews the results of several studies that we have conducted on the nature of speech motor organization. Our primary focus is on intra- and intersegmental organization, but we will also demonstrate the importance of the prosodic component by showing how speaking rate variation shapes segmental organization. These studies investigate three different kinds of data (movement, acoustic, and electromyographic) related to the activities of three different articulators (velum, lips, and tongue). While velar movement data were used by Kent and Minifie (1977) to reveal the inadequacies of earlier theories, we have shown that the patterns they observed conform to the predictions of a model based on the temporal overlap among characteristic movements for adjacent and near-adjacent segments, that is, a coproduction model. The apparent conflicts in the data on both velar and labial coarticulation, previously taken as supporting either the "feature-spread" or the "coproduction" model, are resolved when intrinsic segmental characteristics are considered and the coproduction model is appropriately applied (e.g., Bell-Berti, 1993; Bell-Berti & Krakow, 1991; Boyce, Krakow, & Bell-Berti, 1991; Boyce, Krakow, Bell-Berti, & Gelfer, 1990; Gelfer et al., 1989). EMG data related to tongue activity provide additional support for the segmental interactions of speech that were predicted by the coproduction model. In addition, we have shown that velar function is affected by a variety of nonsegmental factors, including position of the segment within a syllable and within a sentence, as well as by stress and speaking rate (Krakow, 1989, 1993; Krakow, Bell-Berti, & Wang, this volume).[2] Similarly, Krakow (1989, 1993)

---

[2] Furthermore, in stark contrast to previous reports of the velum as a functionally simple articulator, we believe that the segmental and nonsegmental characteristics drawn together here make it clear that the velum is a complex articulator that must be treated as such in any viable model of speech production.

has shown differences between labial articulations for consonants occurring syllable-initially and -finally (see also Browman and Goldstein, this volume, who have shown such differences for lingual articulations), as well as for consonants in stressed and unstressed syllables. These patterns support the notion that the development of a model of speech organization that accounts for the co-occurrence of segmental, syllabic, and prosodic characteristics is imperative.

INTRA- AND INTERSEGMENTAL ORGANIZATION

Intrinsic Segmental Movements

Understanding the nature of intersegmental organization requires an understanding of intrinsic positions for segments, that is, intrasegmental organization. In our work, we have stressed that theories of coarticulation must attend to such information. Data supporting our arguments can be found for a number of different articulators, including the velum (e.g., Bell-Berti, 1980, 1993; Bell-Berti & Krakow, 1991), the lips (e.g., Bell-Berti & Harris, 1979; Boyce, 1988, 1990; Gelfer et al., 1989), and the tongue (e.g., Bell-Berti & Harris, 1974; Harris & Bell-Berti, 1984), and confirmed by other investigators (Marchal, 1988; Silverman & Jun, 1994). Additional demonstrations have been found for the glottis (Munhall & Löfqvist, 1992; Saltzman & Munhall, 1989) and for the pharynx (Parush & Ostry, 1993). These data cover a wide range of measurement techniques, including opto-electrical tracking of lip and velum movement, EMG, acoustic analysis, electropalatography, ultrasound, and transillumination.

Consider the schematic diagrams (Figure 1) that represent two typical velar lowering patterns one is likely to observe during CV_nN sequences (where C represents an obstruent; V_n, some number of vowels; and N, a nasal consonant).[3] Three accounts of these patterns—all widely discussed in the literature—can be shown to fail precisely because they do not consider what each segment, whether phonemically oral or nasal, brings to the sequence.

Moll and Daniloff (1971) applied the feature-spreading model of Henke (1966) to velar movement data. In this account, segments that do not exhibit a contrast for a specific feature simply take on the feature value of the next specified segment in the sequence. Typically, English vowels have been viewed as unspecified with respect to the feature NASAL, as there is no oral-nasal distinction for vowels in English. Under these conditions—that is, assuming vowels to be unspecified for nasality and that Henke's model applies—evidence of velar lowering during the transition between an oral consonant and a vowel, or during a vowel sequence preceding a nasal consonant, is taken as evidence of anticipatory feature spreading from the nasal consonant. Examining the patterns

---

[3] Data reflecting the patterns seen in Figure 1 have been reported elsewhere (e.g., Bell-Berti & Krakow, 1991; Bell-Berti, 1993; Bladon & Al-Bamerni, 1982; Boyce et al., 1990).

presented in Figure 1, the earliest onset of velar lowering in both sequences would be identified as the onset of anticipatory coarticulation for the nasal consonant. Velar lowering would be considered to have spread through the vocalic portions of both utterances. It is important to note that this account does not predict the different patterns shown in the schematic representations of Figures 1(a) and 1(b). Furthermore, this account is largely unidirectional—that is, only the anticipated segment is considered to be important.

In a more recent proposal, Keating (1988a) examined movement patterns like those shown in Figure 1(a). She argued that a pattern in which there is smooth movement between two segments with extreme and opposite positions, despite the occurrence of one or more intermediate segments, should be taken as indicating interpolation through those intermediate segments, and that such segments would then be classified as phonetically unspecified. Although they introduce a timing unit, unspecified segments in Keating's theory contribute nothing to the articulatory trajectory. The notion "unspecified" is shared with Moll and Daniloff (1971), but the nature of coarticulation is different; in Keating's view, the mechanism seems inherently bidirectional (both the start and end points, and the timing of all intervening segments, contribute to the shape of the interpolated trajectory). According to Keating's model, movement patterns in CV_nN sequences like that shown in Figure 1(a) would be taken as evidence of lack of specification of the feature NASAL for English vowels.

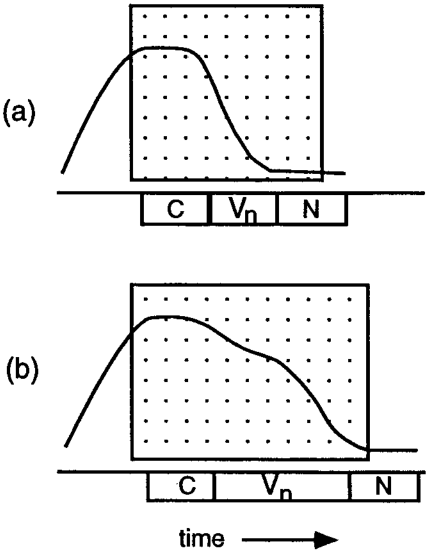

FIGURE 1. Schematic representations of predictions of coproduction model of velar movement patterns for CV_nN sequences.

Keating goes on to say that if a segment shows what looks like a target (i.e., it shows an articulatory position predicted by its featural composition), then it is safe to assume that it enters the motor program with categorical specification (+ or –). In this view, the movement pattern shown in Figure 1(b) would indicate specification of the intervening vowel or vowels with respect to velum position and/or nasalization. However, in a separate paper, Keating (1988b) specifically noted the occurrence of velar movement patterns like that shown in Figure 1(b), where she said that despite being "unspecified," some segments may have limits in the range of articulatory (in this case velar position) variation that they will tolerate; she referred to such limits as "windows." She claimed that this (albeit a wide) spatial window required the velum to remain in an intermediate position, for a period of time, between the extreme high and low positions occurring on either side, giving rise to the pattern presented in Figure 1(b). Keating (1988b) says nothing, however, about when or why one sometimes observes pattern 1(a) and sometimes pattern 1(b) for similar, if not the same, sequences.

Bladon and Al-Bamerni (1982) described the variation in velar movement patterns represented in Figures 1(a) and 1(b), and suggested that speakers simply have the choice between two alternative patterns for the same sequences. Describing pattern 1(a), they said that such movements conformed to the predictions of feature-spreading models, with the time between the onset of velar lowering and the beginning of the nasal murmur being correlated with the duration of the intervening vowel string. Describing pattern 1(b), they said that while the shallow movement onset conformed to the temporal predictions of the feature-spreading models (i.e., it migrates to the beginning of the vocalic portion), the onset of the steeper lowering movement seemed to be in close temporal proximity to the nasal consonant. They offered no way to predict the occurrence of 1(a) vs. 1(b) differentially.

Taken together, the three theories share a lack of insight into what, precisely, phonemically oral segments bring to the articulatory sequence. Fortunately, however, the speech production literature offers considerable insight into the relation between velar height/velopharyngeal port opening and such segmental characteristics as vowel height, and consonant place, manner, and voicing. Velar position is lowest for nasal consonants, being somewhat higher for nasal vowels than nasal consonants (see Henderson, 1984). Furthermore, velar position varies directly with vowel height (Bruckë, 1856; Czermak, 1869; Passavant, 1863; Bell-Berti, Baer, Harris, & Niimi, 1979; Fritzell, 1969; Moll, 1962), and the velum is higher for obstruent consonants than for high vowels (for a detailed discussion of intrasegmental velar organization, see Bell-Berti, 1980, 1993).

An alternative to feature-spread theories is offered by coproduction approach that posits the simple overlap of segmental gestures as the source of most coarticulatory phenomena. We will discuss coproduction in some detail and then show how this approach accounts for the patterns in schematics 1(a) and 1(b). Bell-Berti and Harris (1981), for example, propose that (in the absence of articulatory conflict) the period of time taken up by a given articulator movements (i.e., its "articulatory period") begins at a relatively constant time before the period when the segment dominates the acoustic signal. (Note that this prediction holds even in conditions where articulatory overlap may be expected to increase, and thus affect the measured acoustic period, conditions like faster speaking rates.) The Bell-Berti and Harris theory predicts that an articulator's movements will overlap substantially with the articulatory periods of neighboring segments. In some cases, overlap will extend into the acoustic periods of neighboring segments.[4] This coproduction theory also assumes that intrasegmental organization takes into account the different periods for different articulators so that they are synchronized with respect to each other for the segment's acoustic period. Thus, the timing of articulatory periods bear a constant relation to each other.

A further aspect of the theory is that gestures will combine in an additive fashion for the period of time during which they overlap. Data to this effect were first reported by Bell-Berti (1980), who noticed that when oral consonants in a sequence were tightly overlapped, velar position was often higher than the segments' supposed intrinsic positions might suggest. More extensive work by Boyce (1988) and Munhall and Löfqvist (1992) supports this view of addition as the mechanism for combining simultaneous commands to the articulators. Munhall and Löfqvist (1992) also note that additive combination is frequently found in studies of nonspeech movement control.

In 1991, we tested this model with velar movement data and found that the model predicted the distribution of occurrence of patterns like those in schematics 1(a) and 1(b). We provided clear evidence of a distinct velar movement for the vowel segment (or segments) in CV_nN strings and support for the notion that increasing the duration of the vowel portion will function to separate the overlapping velar lowering gestures for the vowel(s) from that for the nasal consonant (Bell-Berti & Krakow, 1991). Figure 2 presents actual movement data[5] that match the schematic patterns of Figures 1(a) and 1(b). Figures 2(a) and 2(b) show that separate vowel- and nasal consonant-related velar lowering gestures are observed when there is sufficient time between the fricative and the

---

[4] In addition, Bell-Berti and Harris (1981; also Bell-Berti, 1993) suggest that the speech production mechanism adjusts overlap between gestures when the outcome would be unacceptable, and that this adjustment takes the form, when possible, of adjusting the timing of movements (i.e., timing of articulatory periods) to delay their onset. They noted that a delayed onset is more likely to occur when two articulatory gestures that are in conflict occur close together in time. On the other hand, the evidence of the complete movement trajectory is more likely to be seen when two competing gestures are separated by increased temporal intervals.

[5] In the speech samples collected for the Bell-Berti and Krakow (1991) study, /l/ was assumed to be vocalic in nature in terms of velar height because of evidence that it is produced with a velar position more like that of vowels than oral consonants in English (see Kuehn, 1976; Moll & Daniloff, 1971; Ohala, 1971; Schourup, 1973).

nasal, for example, when the vowel sequence is extended by the addition of vocalic segments. Figure 2(c) shows a minimally contrastive sequence without a nasal consonant, from the same data set. Such non-nasal control sequences must be considered in order to distinguish intrinsic velar positions from coarticulatory effects. The data in Figure 2(c) show that the velum lowers for a vowel following a fricative consonant even when there is no nasal consonant in the vicinity. Furthermore, Bell-Berti and Krakow (1991, Figure 8, p. 119) showed that the early part of the velar lowering gesture in CV_nN sequences matches the lowering gesture in minimally contrastive CV_nC sequences—that is, in sequences ending with an obstruent rather than a nasal consonant.

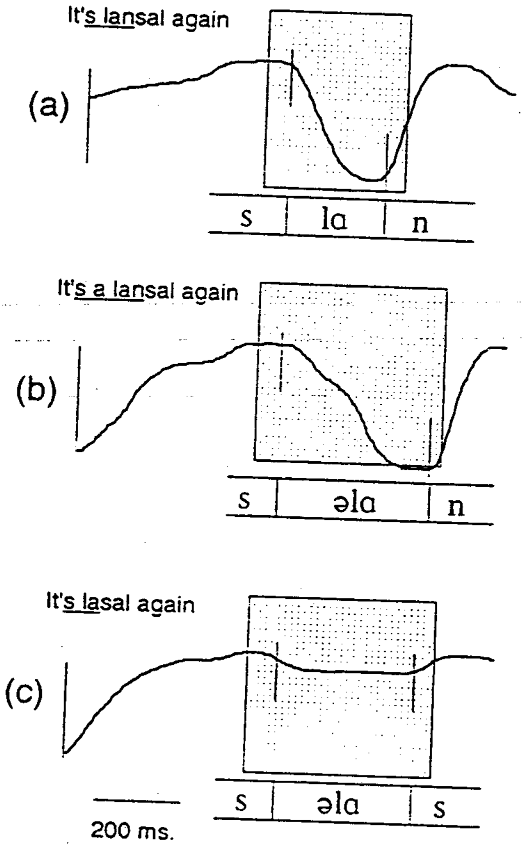

FIGURE 2. Observed anticipatory velar movement patterns for one subject producing "It's lansal" and "It's a lansal;" the relevant segments ([slɑn], [səlɑn], and [səlɑs]) are indicated (adapted from Bell-Berti & Krakow, 1991).

As discussed above, most studies of coarticulation have focused on anticipatory effects, and the examples that we have just provided are limited in the same way. To test the predictions of the coproduction model further, we have recently begun to investigate carryover coarticulation in sequences of the form NV_nC, comparing them to sequences of the form CV_nC. The patterns that we have observed parallel those reported (and described above) for anticipatory coarticulation, lending further support to the model. This work, however, is clearly in its infancy.

Figure 3 provides examples of velar raising patterns when there is sufficient time between the nasal and oral consonants to allow separate vowel- and obstruent-related velar raising gestures to be observed. In the first sequence, [ɪtsmastɪk], the vowel provides a sufficiently long interval (about 300 ms) for the vowel and obstruent gestures to be distinguishable. In the second sequence, [ɪtsmɑlistɪk], separate raising gestures are evident for [l], [i], and [s], leading to the sort of multi-stage pattern shown in Figure 3(b). (In this second sequence, there is a slightly higher position for [ɑ] than [m], a midposition for [l], a high position for [i], and an even higher position for [s]).

As noted above, other types of data confirm the predictions of our model. In what follows, we discuss other aspects of the model and offer evidence from acoustic and articulatory data related to lip movement and then EMG data related to tongue movement.

Articulatory Periods

The studies described above found evidence that individual vowels and consonants have intrinsic velar positioning patterns, that these patterns become visible as time is added to the sequence, and that intrinsic patterns appear to contribute additively to the overall sequence. In a series of studies spanning more than a decade (e.g., Bell-Berti & Harris, 1979, 1982; Boyce, 1988, 1990; Boyce et al., 1991; Boyce et al., 1990; Gelfer et al., 1989), we have examined orbicularis oris EMG activity and lip movement for lip rounding gestures for similar effects, and found confirming evidence for each of these points. In fact, Boyce (1988, 1990) provided evidence that naturally occurring patterns of lip movement for particular words could be synthesized by adding together intrinsic segmental movements extracted from a variety of words with different phonetic contexts. An example of the results of this procedure is shown in Figure 4. Here a synthetic movement trace for the word [kuktluk] has been constructed by adding (ensemble-averaged) movement traces for two other words, [kuktlik] and [kiktluk], that were obtained from a speaker of American English (who produced them in identical carrier phrases). As Figure 4 shows, the "constructed" trace bears a strong resemblance to the true ensemble-averaged [kuktluk] trace.

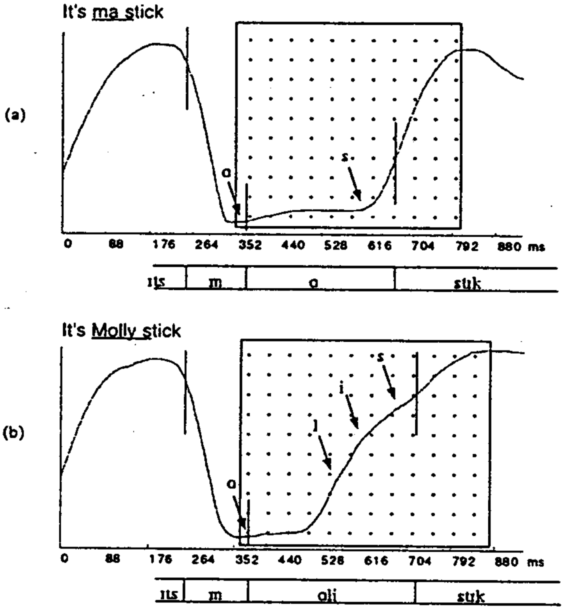

FIGURE 3. Observed carryover velar movement patterns for one subject: (a) producing [mɑ] and (b) [mɑli] utterances; discontinuities assumed to be the result of the onsets of successive segments are indicated by arrows.

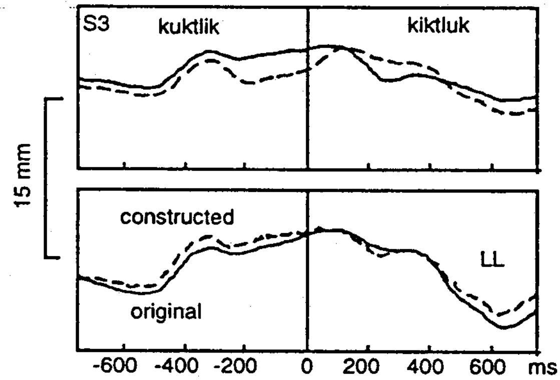

FIGURE 4. Overlaid averaged lip protrusion traces for [kuktlik] and [kiktluk] (top panel); naturally produced averaged lower lip protrusion trace for [kuktluk] with superimposed trace (lower panel) constructed by adding together traces in the upper panel and subtracting averaged trace for [kiktlik]. In all cases, anterior movement is upward (adapted from Boyce, 1990).

One aspect of the model not addressed in our velar and lip studies is whether the acoustic and articulatory periods for a segment have a stable relation in time. We describe here a previously unreported study whose results bear on this issue. Two subjects produced 20 repetitions of each of eight utterance types, comprising four minimal pairs, of the form iC_nV_2, where C_n was [t], [s], [sts], or [stst], and V_2 was [i] or [u]. The sequences were embedded in the carrier phrase, "It's a [__] again." The two subjects were selected from among subjects of our earlier studies (e.g., Bell-Berti & Harris, 1982; Gelfer et al., 1989) because they exhibited different lip activity patterns for [s] and [t]. One subject, S1, used substantial lip activity during [s] and [t] only when the consonants preceded a rounded vowel. This pattern fit the traditional definition of coarticulation from a contextual rounded vowel. If such coarticulation results from overlap of a stable articulatory gesture into the consonant, we might expect that rounding activity to begin at a consistent time. Further, the intrinsic rounding specification for the consonant is "unrounded" (i.e., no intrinsic rounding movement or intrinsic lip-spreading movement). The second subject, S2, produced alveolar consonants with substantial labial activity regardless of the phonetic properties of the following vowel. Such a pattern suggests that for this subject, the intrinsic rounding specification for alveolar consonants is "rounded." Again, we expect rounding activity to begin at a consistent time, but that time should be earlier in the case of S2.

We measured the frequency of resonances in the region of F2 from the beginning of V2 (the release burst of [t] or the end of frication of [s]), at 40 ms intervals back through the consonant sequence to the end of the first vowel, [i].[6] In all cases, the earliest measurement was made at the end of the first vowel. We took the ensemble average of the frequency measures at each time interval (the average of the 20 repetitions of each type, separately for each subject), and used t-tests to compare the significance of the differences between averages for the members of each minimal pair (i.e., [iC_ni] vs. [iC_nu]) at each interval.

Since lip protrusion should have the effect of lowering vocal tract resonances (Fant, 1960; Stevens & House, 1961), anticipation of lip rounding should result in lower frequency F2 resonances in consonants preceding a rounded vowel than in consonants preceding a nonrounded vowel. Since [s] and [t] share the same place of articulation, the acoustical resonances for these segments should be similar in the same phonetic contexts. Based on the observation that S2 does not differentiate lip activity for consonants preceding [i] and [u], we expected the two subjects to differ in the effect of the upcoming vowel on resonances in the frication. In particular, we expected S1 to show higher resonances before [i].

As expected, for both subjects there were significant differences at the onset of the second vowel for all utterance pairs, although this difference was larger for S1; these acoustic data are plotted in Figure 5. In addition, S1 showed a

---

[6] No measurements were made of a 40-ms point if it fell during a [t] closure.

significant effect in the predicted direction only for /isV/, and only for the frication within 80 ms of V2. On the other hand, S2 showed significant differences in the predicted direction for /isV/ and /istsV/ sequences, with the differences occurring as early as 160 ms before the second vowel. S2 also showed a significant difference in the predicted direction in the /iststV/ sequence, but only as early as 80 ms before V2. For both subjects, there were also scattered and inexplicable differences in the direction opposite to that predicted; that is, where resonances preceding [u] were higher than those preceding [i]. We would conclude that overall these acoustic data reflect the coproduction of the vowels [i] and [u] with the preceding consonant string. Specifically, the coarticulatory effects (i.e., resonances preceding [u] being lower than those preceding [i]) began no more than 160 ms before V2, even when the consonant sequences were more than twice that duration.

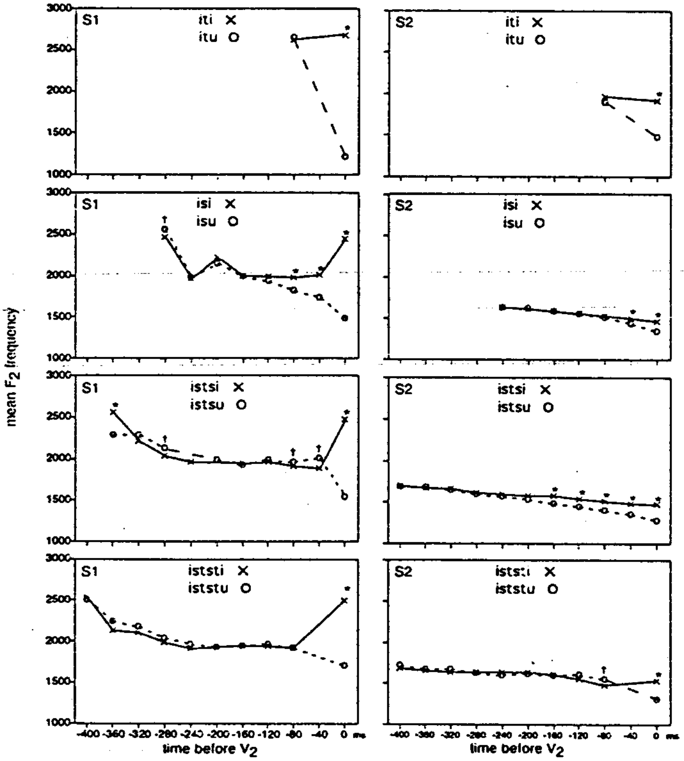

FIGURE 5. Ensemble-averaged F_2 frequency plots for two subjects (S1 and S2) for four minimally contrastive [iC$_n$i] and iC$_n$u] utterances; F_2 measurements occur at 40 ms intervals. Significantly different means are marked with '*' for comparisons in which the means of V_2=[i] were higher, and '†' for comparisons in which the means of V_2=[u] were higher.

In their 1982 paper, Bell-Berti and Harris reported not only that the onset of orbicularis oris EMG activity for [u] showed limited anticipatory effects, but also that the offset of orbicularis oris activity for the first vowel in [uC_ni] and [uC_nu] sequences did not extend any further into longer sequences than shorter ones (i.e., carryover coarticulation was temporally limited). Furthermore, they also reported that lip position for the following vowel did not influence the timing of the end of the first vowel gesture. Hence, these early EMG labial data complement the velar carryover data described above. They also reflect the pattern, described immediately below, in which articulatory activity for the vowels in a VCV sequence is suppressed during the consonantal articulation.

Articulatory Overlap

Among the earliest studies in this series were some that called into question the feature-spread models in general, and Henke's (1966) lookahead model in particular. In 1974, Bell-Berti and Harris reported EMG recordings from the genioglossus muscle, which is active for the forward, bunching and raising gesture for [i]. That study examined data for three speakers who produced a series of four-syllable nonsense utterances of the form [əpipiCə], where C was [p] or [b] and stress was systematically varied between the second and third syllables. The EMG signals were ensemble averaged, using the moment of closure for the second [p] as the reference point for aligning the signals of the 15-20 tokens of each of the eight utterance types. For the subject whose data are presented in Figure 6, as for the other two, the data are striking for the biphasic nature of the genioglossus activity for each utterance type, with one burst for the first [i] and a separate burst for the second [i]. Furthermore, closer inspection reveals that, for each minimal comparison, the duration of the interval between bursts was longer and the ensemble-average EMG minimum between the vowels was lower when stress fell on the second [i]. As one might expect, the acoustic duration of the medial [p] closure was longer when it immediately preceded the stressed vowel. This fits with what we know about the effects of stress on consonants in prestressed position: acoustic silent durations are longer, bursts are stronger. Moreover, movement displacements are greater; and stressed vowels are longer. The effect of these differences is to alter the temporal separation of the [i] gestures, and, as a consequence, their overlap is decreased. Or, in more current usage, we would say that there was less coproduction of the [i] gestures of adjacent syllables, and less overlap means more time for the tongue to relax from its [i]-related position. Thus, the troughs in the EMG traces are lower when the second vowel is stressed.

Additional confirmation of this pattern can be found in Boyce (1988), who found that double-peaked "trough" patterns occurred for lip movement and EMG orbicularis oris activity for sequences like [kuktluk], [kuktuk], and [kukuk], and that the troughs became shallower (i.e., the lip displacement or EMG activity minimum between peaks decreased) when fewer consonants intervened between rounded vowels. This effect is shown in Figure 7, which shows ensemble-averaged lip movement (S1, above) and EMG activity (S3, below) traces for the three words, spoken by one male (S1) and one female (S3) speaker of American English. Statistical significance of this effect was established by comparing coefficients of quadratic trend in peak-to-peak portions of the traces.

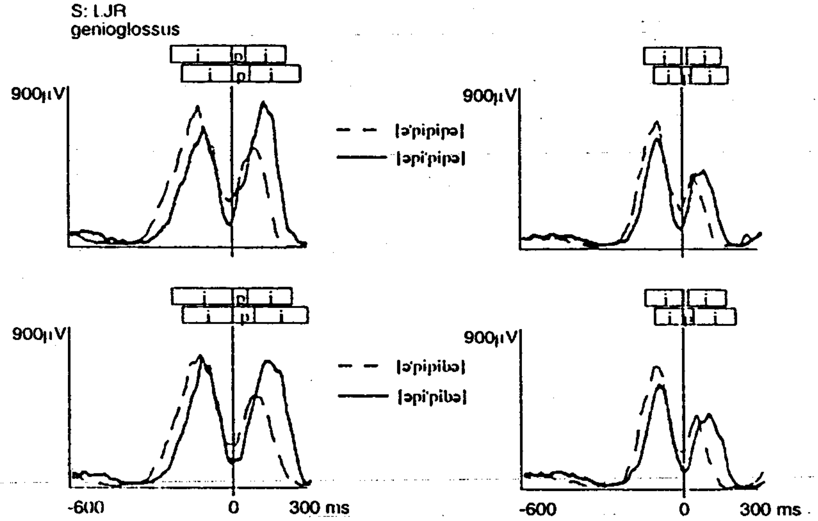

FIGURE 6. Ensemble-average genioglossus EMG activity for one speaker producing four minimally contrastive utterances at normal rate (left-hand panels) and four at fast rate (right-hand panels) (adapted from Bell-Berti & Harris, 1974).

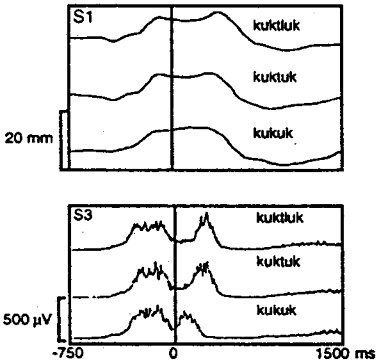

FIGURE 7. Ensemble-averaged lower-lip protrusion movement for [kuktluk], [kuktuk], and [kukuk], produced by S1 (above), anterior movement is upward; ensemble-averaged orbicularis oris EMG activity for [kuktluk], [kuktuk], and [kukuk], produced by S3 (below), (adapted from Boyce, 1988).

SUMMARY

Observations of multistage velar lowering before nasal consonants (anticipation) and multistage velar raising after nasal consonants (carryover) are, we believe, equivalent to the observation of lingual EMG and displacement minima, or "troughs," for [iCi] sequences (e.g., Bell-Berti & Harris, 1974; Harris and Bell-Berti, 1984) and of labial EMG and displacement minima for [uC$_n$u] sequences (e.g., Bell-Berti & Harris, 1982; Boyce, 1988; 1990). That is, these troughs, along with the stable time intervals found for the beginnings of segmental gestures (when appropriate control utterances are used), are reflections of the unitary nature of segments. Taken together, these data on velar, labial, and lingual articulation all share a common characteristic: they reflect segment-by-segment organization within overlapping temporal windows.

It is also important to examine the temporal cohesion of the gestures that comprise a segment, since an important assumption of the Bell-Berti and Harris (1981) model is the synchronization of a segment's component gestures. Happily, support for the idea of tightly constrained within-segment timing of component articulatory gestures can be found in research on the effects of perturbations of one articulator on the relative timing of gestures of another (see, for example, Saltzman, Löfqvist, Kinsella-Shaw, Kay, & Rubin, this volume).

PROSODIC ORGANIZATION: SPEAKING RATE

Because there are other sources of influence on the resulting movement ensemble, identification of a segmental level of gestural organization and an understanding of the manner in which gestures for successive segments combine provide only a partial understanding of the nature of speech motor organization. These influences include, but are not limited to, variations in speaking rate, syllable structure, stress, and the location of a syllable within a phrase or sentence. Of these, we have selected the effects of speaking rate on velar movements for discussion in this chapter.

The data on rate variation shed light on precisely the sorts of patterns we have just been addressing, and on the nature of variation between the patterns shown in the schematics in Figures 1(a) and 1(b). The coproduction model predicts that when speech rate is increased, there will be increased overlap of gestures for adjacent and near-adjacent segments. Hence, increasing the speaking rate in the production of a CV$_n$N sequence ought to have an effect that resembles that of decreasing the number of intervening vowels—that is, a smoother lowering movement ought to be evident. Figure 8(a) shows this effect on the movement patterns of a speaker producing the same utterance at two different rates (adapted from Bell-Berti & Krakow, 1991). We would also expect similar variation to occur when two different speakers produce the same sequence at different rates, and Figure 8(b) shows movement patterns of two speakers, one whose "comfortable" rate was much faster than the "comfortable" rate of the other (adapted from Bell-Berti and Krakow, 1991).

Other studies on the effects of speaking rate show a reduction in the magnitude of velar movements at faster rates (e.g., Kent, Carney, & Severeid, 1974; Kuehn, 1976). Both types of observation (i.e., a decrease in multi-stage gestures and a reduction in positional extremes) are predicted by the coproduction model, as both are outcomes of increased temporal overlap of gestures, although studies have generally examined only one or the other effect. Subsequent to the publication of Bell-Berti and Krakow (1991), we reexamined the data and measured the height of the velum at the release of the /s/ of "It's" or "It's say" for each sequence containing a nasal consonant. The results revealed that the velar peak was consistently higher in the normal than the fast rate productions, indicating a reduction in overlap between the gesture for the oral consonant and that for the nasal consonant at the normal rate. Still further support for this interpretation of these data can be found in the fact that the velar peak was higher when more vocalic segments intervened between the /s/ and the /n/ as shown in Figure 9.

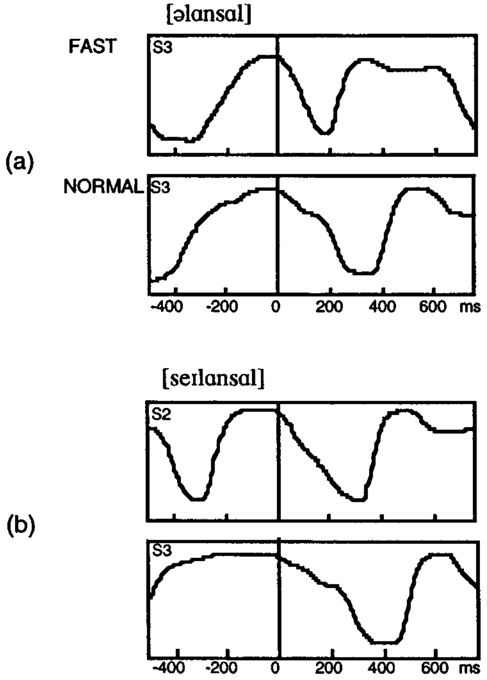

FIGURE 8. Velar movement patterns for (a) one subject producing "It's a lansal" at her "normal and "fast" speaking rates; (b) two subjects producing "It's say lansal" at their normal speaking rates (adapted from Bell-Berti & Krakow, 1991).

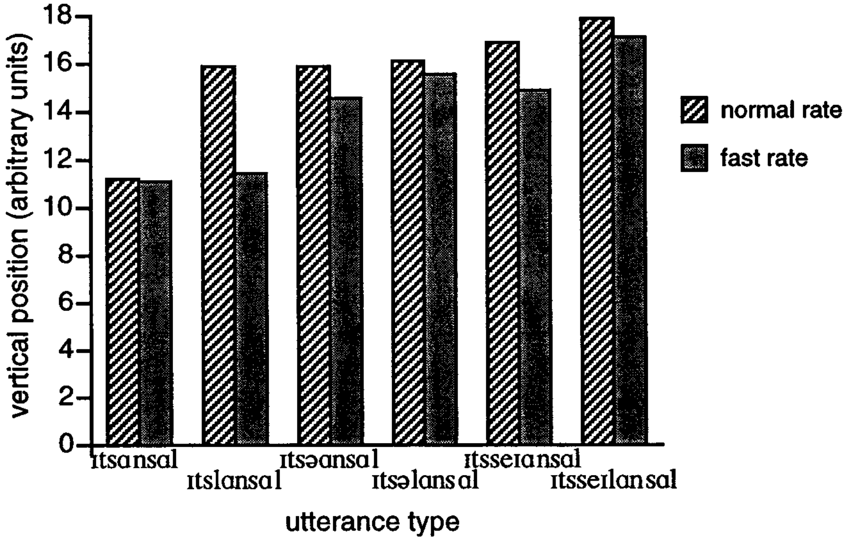

FIGURE 9. Mean peak velar position during [s] of "It's" or "It's say" in each of the six utterances shown at the bottom. Each mean represents 12 tokens produced by one subject at each of two rates (adapted from Krakow, 1993).

These results show us how to predict when continuous single-stage velar lowering will occur and when complex, multi-stage velar lowering will occur. These results also make it possible to reconcile the data reported by Bladon and Al-Bamerni (1982) and by Keating (1988a, b) with those reported by Bell-Berti and Krakow (1991), and to refute the conflicting interpretations of Bladon and Al-Bamerni (1982) and of Keating (1988a, b) in favor of the coproduction model.

CONCLUSION

Drawing upon studies of three articulators, studies that employed very different types of measurements, we have found that there is actually rather little "anticipation" of articulatory activity. Whether in the "multi-stage" gestures observed in velar movement patterns, or the failure to find extended acoustical differences for consonant sequences preceding rounded and nonrounded vowels, or the adjustment of gesture onset timing to acoustic output constraints on lingual and labial EMG and displacement (corresponding to the duration of the intervening consonant occlusions), segmental articulatory behaviors do not extend very far from the segments for which they were intended. That is, all of the data discussed here, and much of the data presented in the literature, whether displacement, acoustical, or EMG data, support the predictions of the coproduction model.

That this is true for the extensive data on anticipatory coarticulation seems fairly obvious; we hold that the, thus far, rather limited data on carryover coarticulation are also consistent with a coproduction interpretation. The preliminary carryover data on velar function presented here appear to reflect the same time-stable patterns as those reported by Bell-Berti and Harris (1982) for lip-rounding gestures. Clearly, though, much additional work needs to be done to describe adequately the time course of segment offsets and their effects on nearby segments.

We have also examined the interaction between one prosodic effect, speaking rate, with intra- and intersegmental behaviors. Clearly, one effect of increased speaking rate is increased temporal overlap of the gestures for successive segments. But additional data are needed before we are able to provide a model that integrates segmental, syllabic, and prosodic factors to account for the range of observed coarticulatory effects.

ACKNOWLEDGMENT

The research reported here and the preparation of this manuscript were supported by NIDCD grant DC-00121 to the Haskins Laboratories.

REFERENCES

Beckman, M., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty & D. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, tone* (pp. 68-86). Cambridge: Cambridge University Press.

Bell-Berti, F. (1980). A spatial-temporal model of velopharyngeal function. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice, Volume IV* (pp. 291-316). New York: Academic Press.

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental content. In M. K., Huffman & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (*Phonetics and Phonology,* Vol. 5, pp. 63-85). New York: Academic Press.

Bell-Berti, F., & Harris, K. S. (1974). More on the motor organization of speech gestures. *Haskins Laboratories Status Report on Speech Research, SR-37/38,* 73-77.

Bell-Berti, F., & Harris, K. S. (1979). Anticipatory coarticulation: Some implications from a study of lip rounding. *Journal of the Acoustical Society of America, 65,* 1268-1270.

Bell-Berti, F., & Harris, K. S. (1981). A temporal model of speech production. *Phonetica, 38,* 9-20.

Bell-Berti, F., & Harris, K. S. (1982). Temporal patterns of coarticulation: Lip rounding. *Journal of the Acoustical Society of America, 71,* 449-454.

Bell-Berti, F., & Krakow, R. A. (1991). Anticipatory velar lowering: A coproduction account. *Journal of the Acoustical Society of America, 90,* 112-123.

Bell-Berti, F., Baer, T., Harris, K. S., & Niimi, S. (1979). Coarticulatory effects of vowel quality on velar function. *Phonetica, 36,* 187-193.

Bladon, R. A. W., & Al-Bamerni, A. (1982). One-stage and two-stage temporal patterns of velar coarticulation. *Journal of the Acoustical Society of America,* 72, S104. (A)

Boyce, S. E. (1988). *The influence of phonological structure on articulatory organization in Turkish and in English.* Unpublished doctoral dissertation, Yale University.

Boyce, S. E. (1990). Coarticulatory organization for lip rounding in Turkish and in English. *Journal of the Acoustical Society of America, 88,* 2584-2595.

Boyce, S. E., Krakow, R. A., & Bell-Berti, F. (1991). Phonological underspecification and speech motor organization. *Phonology, 8,* 219-236.

Boyce, S. E., Krakow, R. A., & Bell-Berti, F., & Gelfer, C. E. (1990). Converging sources of evidence for dissecting articulatory movements into gestures. *Journal of Phonetics, 18,* 173-188.

Browman, C. P., & Goldstein, L. (1991). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 341-376). Cambridge: Cambridge University Press.

Bruckë, E. W. (1856). *Grundzuge der Physiologie und Systematik der Sprachlaute.* Vienna: Carl Gerold's Sohn (cited in Fritzell, 1969).

Czermak, J. N. (1869). Wesen und Bildung der Stimmund Sprachlaute. In *Czermak's gesammelte Schriften, Vol. 2.* Leipzig: Wilhelm Engelman (cited in Fritzell, 1969).

Daniloff, R. G., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics, 1,* 239-248.

deJong, K. (1991). An articulatory study of consonant-induced vowel duration changes in English. *Phonetica, 48,* 1-17.

Fant, C. G. M. (1960). *Acoustic theory of speech production.* The Hague: Mouton.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing control. *Journal of Phonetics, 8,* 113-133.

Fowler, C. A. (1984). Current perspectives on language and speech production: A critical review. In R. Daniloff (Ed.), *Recent advances in speech, hearing, and language, Vol. 3* (pp. 195-278). Boston: College-Hill Press.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36,* 171-195.

Fritzell, B. (1969). The velopharyngeal muscles in speech: An electromyographic-cinéradiographic study. *Acta Otolaryngologica, Supplement 250.*

Fujimura, O. (1990). Methods and goals of speech production research. *Language and Speech, 33,* 195-258.

Gelfer, C. E., Bell-Berti, F., & Harris, K. S. (1989). Determining the extent of anticipatory coarticulation: Effects of experimental design. *Journal of the Acoustical Society of America, 86,* 2443-2445.

Harris, K. S., & Bell-Berti, F. (1984).On consonants and syllable boundaries. In L. J. Raphael, C. B. Raphael, & M. R. Valdovinos (Eds.), *Language and cognition* (pp. 89-95). New York: Plenum Publishing.

Henderson, J. B. (1984). *Velopharyngeal function in oral and nasal vowels: A cross-language study.* Unpublished doctoral dissertation, University of Connecticut, Storrs.

Henke, W. (1966). *Dynamic articulatory model of speech production using computer simulation.* Unpublished doctoral dissertation, MIT.

Keating, P. A. (1988a). Underspecification in phonetics. *Phonology, 5,* 275-292.

Keating, P. A. (1988b). The window model of coarticulation: Articulatory evidence. *UCLA Working Papers in Phonetics, 69,* 3-29.

Kent, R. D., Carney, P. J., & Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model of binary control. *Journal of Speech and Hearing Research, 17,* 470-488.

Kent, R. D., & Minifie, F. D. (1977). Coarticulation in recent speech production models. *Journal of Phonetics, 5,* 115-133.

Krakow, R. A. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures.* Unpublished doctoral dissertation, Yale University, New Haven, CT.

Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate. In M. K., Huffman & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum (Phonetics and Phonology,* Vol. 5, pp. 87-114). New York: Academic Press.

Kuehn, D. P. (1976). A cineradiographic investigation of velar movement in two normals. *Cleft Palate Journal, 13,* 88-103.

Lindblom, B. E. F. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America, 35,* 1773-1781.

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

Marchal, A. (1988). Coproduction: Evidence from EPG data. *Speech Communication, 7,* 287-295.

Moll, K. (1962). Velopharyngeal closure in vowels. *Journal of Speech and Hearing Research, 5,* 30-77.

Moll, K. L., & Daniloff, R. G. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America, 50,* 678-684.

Munhall, K. G., & Löfqvist, A. (1992). Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics, 20,* 111-126.

Nittrouer, S., Munhall, K., Kelso, J. A. S., Tuller, B., & Harris, K. S. (1988). Patterns of interarticulator phasing and their relation to linguistic structure. *Journal of the Acoustical Society of America, 84,* 1653-1661.

Parush, A., & Ostry, D. J. (1993). Lower pharyngeal wall coarticulation in VCV syllables. *Journal of the Acoustical Society of America, 92,* 715-722.

Passavant, G. (1863). Ueber die Verschliessung des Schlundes beim Sprechen. Frankfurt a.M.: J. D. Sauerlander (cited in Fritzell, 1969).

Silverman, D., & Jun, J. (1994). Aerodynamic evidence for articulatory overlap in Korean. *Phonetica, 51,* 210-220.

Saltzman, E., & Munhall, K. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1,* 333-382.

Stevens, K. N., & House, A. S. (1961).An acoustical theory of vowel production and some of its implications. *Journal of Speech and Hearing Research, 4,* 303-320.

Turk, A. (1994). Articulatory-phonetic clues to syllable affiliation: Gestural characteristics of labial stops. In P. Keating (Ed.), *Papers in laboratory phonology III: Phonological structure and phonetic form* (pp. 107-135). Cambridge: Cambridge University Press.

Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica, 45,* 122-139.