Anticipatory velar lowering: A coproduction account[a)]

157

Fredericka Bell-Berti
Department of Speech, Communication Sciences, and Theatre, St. John's University, Jamaica, New York 11439 and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

Rena Arens Krakow
Department of Speech-Language-Hearing, Temple University, Philadelphia, Pennsylvania 19122 and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

(Received 17 August 1990; accepted for publication 11 March 1991)

Feature spreading and coproduction models make fundamentally different assumptions about the nature and organization of speech motor control, and yet each model is supported by some, but not all, of the existing empirical data. This has led some researchers to conclude that speakers probably use alternative strategies at different times. This study suggests that the identification of coarticulatory influences requires the concurrent identification of intrinsic articulatory characteristics of the segment. Moreover, the evidence for feature spreading or variable coarticulation strategies derives from the misidentification of such intrinsic characteristics as context effects. This velar coarticulation study used a controlled comparison between CV_nN and CV_nC minimal pairs, where C is an oral consonant, V_n is any number of vowels, and N is a nasal consonant. Vocalic string duration was manipulated by varying the number of segments and speech rate, allowing us to alter the time between the onsets of vocalic and subsequent consonantal gestures. Velar lowering occurred in CV_n sequences, whether or not a nasal consonant followed, and similar vocalic gestures were observed across minimally contrastive environments with and without the nasal consonant. Moreover, velar lowering for the nasal consonant began in close temporal proximity to the nasal murmur. These results strongly support the coproduction model and provide insight into previously conflicting reports.

PACS numbers: 43.70.Aj, 43.70.Bk, 43.70.Jt

INTRODUCTION

In early studies of speech production, coarticulation was viewed as problematic because researchers' intuitions about the speech stream (the notion of "segments") did not match the acoustic or articulatory output that the new technology revealed (Harris, 1970; Kent and Minifie, 1977; Liberman *et al.*, 1962). To continue to maintain the notion that speech is, at some level, segmentally structured, 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. Over the years, two such general types of explanations of observed coarticulatory phenomena have been offered: "feature spreading" models (e.g., Henke, 1966; Joos, 1948; Kozhevnikov and Chistovich, 1965; Moll and Daniloff, 1971) and "coproduction" models (e.g., Bell-Berti and Harris, 1981; Browman and Goldstein, 1986; Fowler, 1980; Öhman, 1966).

Studies of anticipatory nasal coarticulation (including data on velar lowering movements, nasal airflow, velopharyngeal port opening, or levator palatini relaxation) along with studies of labial coarticulation (including data on upper lip protrusion, lower lip protrusion, or orbicularis oris contraction) have been widely cited in studies proposing the different coarticulation models (e.g., Bell-Berti and Harris, 1982; Benguerel and Cowan, 1974; Daniloff and Moll, 1968; Kent *et al.*, 1974; Moll and Daniloff, 1971). The present study reexamines the theoretical and empirical bases for the two models with a focus on nasal coarticulation. (We refer the reader to similar discussions of lip rounding in Boyce, 1988; Boyce *et al.*, 1990; Gelfer *et al.*, 1989.) We begin by reviewing the models themselves.

In "feature spreading" models, an articulatory planning (or "look-ahead") component determines what movements are required for upcoming segments and initiates them as soon as their onset would not interfere with the more immediate articulatory requirements [Fig. 1(a)]. In these theories, the temporal extent of such anticipatory adjustments is limited only by characteristics of other segments (e.g., Henke, 1966), or of the syllabic (e.g., Kozhevnikov and Chistovich, 1965) or syntactic structure of the utterance (e.g., McClean, 1973). In the absence of such context-conditioned limitations, however, these theories claim that coarticulation may be unlimited in extent. Thus, for example, velar lowering for a nasal consonant is predicted to spread back to the first vowel in a CV_nN sequence,[1] but not to go beyond the first vowel because of the conflicting high velum specification of the initial oral consonant. This is schematized in Fig. 1(a), where the onset of velar lowering for the nasal consonant is indicated with a solid vertical line: this onset

a) Portions of this paper were presented at the Convention of the American Speech-Language-Hearing Association, Boston, MA, November 1988, and at the meeting of the Linguistic Society of America, Washington, DC, December 1989.

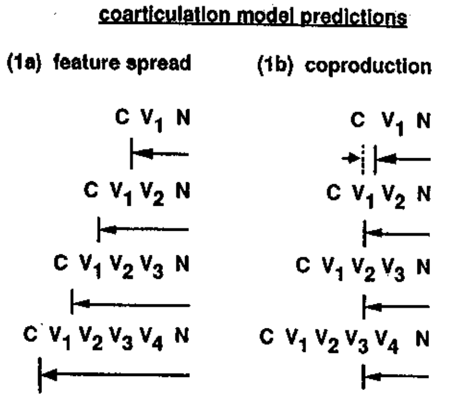

FIG. 1. Predictions of feature spread and coproduction models. (a) Feature-spread models predict that velar lowering in anticipation of a nasal consonant extends to the beginning of the vocalic sequence preceding the nasal consonant, regardless of vocalic string duration or number of segments. (b) Coproduction models predict that velar lowering during a vocalic sequence preceding a nasal consonant begins at a stable time before the nasal consonant, regardless of vocalic string duration.

occurs earlier before the nasal consonant as the vowel sequence becomes longer, being limited only by the presence of the oral consonant.

In contrast, "coproduction" models claim that coarticulation is a function of the overlap of gestures whose onsets bear a stable relation to other aspects of the articulation of a given segment [Fig. 1(b)]. For example, Bell-Berti (1980) has proposed that the onset of velar lowering for a nasal consonant (or velar raising for an obstruent consonant) bears a stable temporal relation to the achievement of the oral tract constriction that is characterized in units of time, not numbers of segments. Furthermore, empirical evidence, from studies of the velum and from the lips, in support of these models suggests that gestural onsets are of limited temporal extent (e.g., Bell-Berti, 1980; Bell-Berti and Harris, 1982; Gelfer *et al.*, 1989). In this type of model, there is no "look-ahead" component that alters the plan for a given segment as a function of context. That is, in coproduction models coarticulation does not affect the nature of the individual gestures; rather, it is manifested in the way in which co-occurring and successive stable gestures combine (Bell-Berti and Harris, 1981; Munhall and Lofquist, 1991; Saltzman and Munhall, 1989). In this model, gestures are modified only by having their onsets delayed by the requirements of preceding segments (i.e., carryover coarticulation), exemplified in the first row of Fig. 1(b), where the dashed vertical line indicates the time at which velar lowering for the nasal consonant would have begun if its onset had not been delayed by the requirements for the initial oral consonant (that is, the time between the dashed and solid lines represents a period of carryover coarticulation).

Clearly, the two types of models, coproduction and feature spreading, make fundamentally different assumptions about the nature and organization of speech motor control. However, each model appears to account for some, but not all, of the data. Thus we are confronted with a new problem of coarticulation. That is, we are left either to accept the unparsimonious explanation that production strategies may vary according to articulator, speaker, and token, or to consider the possibility that some of the data have been interpreted inappropriately. We will, in fact, argue that one of the major failures within studies of coarticulation, on the part of both feature spreading and coproduction theorists, has been the assumption that one can identify coarticulatory influences of a segment's context without first identifying intrinsic articulatory characteristics of the segment itself. We now turn to our explanation of this notion with respect to nasal coarticulation.

In spite of some evidence to the contrary, a general assumption in many studies of anticipatory velar lowering has been that the articulatory plan calls for a uniformly low velar position for nasal consonants, a uniformly high velar position for oral consonants, and a uniformly neutral position for vowels. According to this view, for anticipatory coarticulation to occur, the neutral specification of the vowel must be replaced, through feature copying, with the next specified velar position. Data from a number of studies showing that the velum begins to lower at the consonant release in CV_nN strings have been taken to support feature spreading models (e.g., Kent *et al.*, 1974; Moll and Daniloff, 1971; Ohala, 1971).

In contrast, Bell-Berti reported data explicitly calling into question the uniformity, with respect to velar height, of each of the three categories of sounds that have been described, nasal consonants, oral consonants, and vowels. In one study, she showed variable velar positions for different vowels as a function of vowel height, as well as effects of vowel height on velar position for adjacent oral or nasal consonants (Bell-Berti *et al.*, 1979). In a subsequent study, Bell-Berti (1980) showed that, within a given vowel environment, velar raising continued through long oral consonant sequences (of 400 ms or so), achieving the highest positions for the longest sequences. These data suggest an additive effect from the overlapping contributions of intrinsic velar positions for adjacent oral consonants. The results also showed that the effect of the following vowel on velar position during the consonant sequence extended back only about 250 ms; that is, the vowel affected velar position during the preconsonantal vowel when the consonant sequence was very short, but had no effect on the preconsonantal vowel or the early part of the consonant sequence when the consonant sequence was very long. This suggests a limited temporal window for coarticulatory effects of velar position, and finds support in the work of Ushijima and Hirose (1974) on Japanese and Benguerel *et al.* (1977) on French. Furthermore, whereas Bell-Berti's results are incompatible with feature-spreading models, they are predicted by coproduction accounts of speech production.

In an attempt to sort out the two general types of models, Bladon and Al-Bamerni (1982) concluded that both means of coarticulation are employed by individual speakers. Thus they reported that, in some instances for each speaker, velar lowering occurred at variable times before a nasal consonant correlating with the duration of the vocalic string. In other instances, velar lowering gestures showed a "two-stage" movement, with the first stage occurring well in advance of the nasal consonant, in accord with feature spreading models, and the second stage occurring as though time locked to the onset of the closure for the nasal consonant, in accord with coproduction models. Based on these results, Bladon and Al-Bamerni suggest a model of speech production that allows variable production strategies. Similar results in studies of lip protrusion have led Perkell to propose a "Hybrid" model (Perkell, 1986; see also Perkell and Chiang, 1986). Neither Perkell nor Bladon and Al-Bamerni proposed a specific source for a speaker's choice among variable patterns.

In the present study, we will suggest that much, if not all, of the evidence for feature spreading (whether in variable strategy accounts or in straight feature spreading accounts), may be due to the mistaken identification of an intrinsic articulatory component of a given segment for a coarticulatory effect of an adjacent or near-adjacent segment (Bell-Berti, 1980). In part, the mistake is due to the adoption of phonologically based characterizations in some models of speech motor control. For example, with respect to velar activity and nasalization, it has seemed appropriate to assume that, for English, phonological description needs to be concerned only with the specification [+] or [−] NASAL for consonants, but that it need not be concerned with such a specification for vowels, since vowel nasalization is not distinctive. In the articulatory domain, however, we must also be concerned with documented differences in velar position for different vowel phonemes in different languages (e.g., Bell-Berti, 1980; Bell-Berti *et al.*, 1979; Fritzell, 1969; Henderson, 1984; Moll, 1962; Passavant, 1963; Ushijima and Sawashima, 1972).[2] Moreover, velar height for oral vowels has been shown to be lower than that for oral consonants. Thus, when the coarticulation accounts describe the earliest onset of velar lowering in a CV_nN sequence as the onset of coarticulation, they may, in fact, only have identified an expected transitional movement of the CV sequence, a movement that is unrelated to the nasal consonant. To argue otherwise requires clear evidence that such movements occur only in the context of the upcoming nasal.

This study, in contrast to previous studies of velar coarticulation, included six minimally contrastive CV_n sequences followed by either an oral or a nasal consonant (i.e., CV_nC vs CV_nN). The sequences were designed to allow us to systematically distinguish coarticulatory effects of the nasal consonant from intrinsic articulatory gestures of the adjacent segments. In addition, we adopted the strategy used in most earlier studies of lengthening the vocalic sequence by increasing the number of segments (up to three) to examine the extent of anticipatory nasal coarticulation after factoring out the intrinsic effects. We also used changes in speaking rate as an alternative way of manipulating sequence duration, also allowing us to explore the effects of suprasegmental manipulation on the extent of coarticulation. We argue here that the coproduction account receives strong support when intrinsic segmental effects are distinguished from coarticulatory effects.

I. METHODS

A. Subjects

The subjects were three native speakers of dialects of American English that are spoken in the Metropolitan New York area. None of the subjects reported a history of speech or hearing disorder. Subject 1 is a coauthor of the paper.

B. Speech samples

Table I lists the target sequences for this study: half of the utterances, the "nasal utterances," contained a post-vocalic nasal consonant, /n/; the other half, the "oral utterances," contained a postvocalic oral consonant, /s/. The target words varied in the number of vocalic segments (from one to three) preceding the consonant. The vocalic segment immediately preceding the nasal or oral consonant was /ɑ/;/ɑ/ was chosen to maximize velar lowering. Each target was embedded in a carrier phrase, "It's____again." According to the feature spreading models, the longer the vocalic string, the earlier we can expect to see the onset of velar lowering for the nasal consonant. Effects of sequence length were examined by comparing sequences with different numbers of vocalic segments and also by comparing sequences produced at different speech rates. Given the matched sequences without nasal consonants, it was possible to determine at what point in the vocalic string velar lowering for the nasal consonant was evident as distinct from velar lowering for the vocalic string itself. Note that we are taking /l/ 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, for example, Kuehn, 1976; Moll and Daniloff, 1971; Ohala, 1971; Schourup, 1973).

The test words were presented to the subjects on individual index cards. Subjects were asked to produce six repetitions of each utterance in each of two test orders, for a total of 12 tokens of each of the 12 utterance types per subject. The subjects were asked to produce the items at a self-selected "conversational" speaking rate. To allow us to examine the relation between speech rate and gestural organization, subject 3 was asked to produce an additional 12 repetitions of the list at a rapid rate. Because subject 2's "conversational" rate was rapid, we had data from two subjects (subjects 1 and 3) at a slower rate, and from two subjects (subjects 2 and 3) at faster rates (Fig. 2). Subjects made a small number of errors and, in a few instances, produced additional repetitions, so that there were occasionally fewer and occasionally

TABLE I. Experimental utterance set.

Oral		Nasal	
/ɑsɑl/	(asal)	/ɑnsɑl/	(ansal)
/lɑsɑl/	(lasal)	/lɑnsɑl/	(lansal)
/ʌ ɑsɑl/	(a asal)[a]	/ʌ ɑnsɑl/	(a ansal)[a]
/ʌ lɑsɑl/	(a lasal)	/ʌ lɑnsɑl/	(a lansal)
/seɪ ɑsɑl/	(say asal)	/seɪ ɑnsɑl/	(say ansal)
/seɪ lɑsɑl/	(say lasal)	/seɪ lɑnsɑl/	(say lansal)

[a] Although these are not standard English sequences, subjects were able to produce these sequences as requested, without using "an" as the indefinite article.

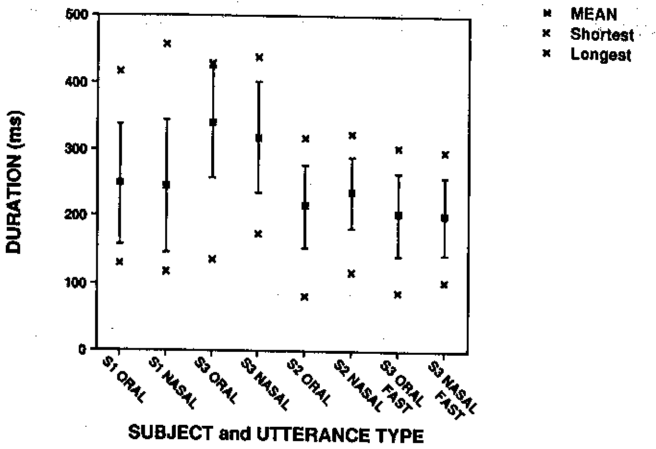

FIG. 2. Mean vocalic sequence durations, standard deviations, and ranges (in ms) of oral and nasal utterances at the self-selected normal rate for all three subjects and at the fast rate for subject 3.

more than 12 tokens of each utterance type. Thus altogether we recorded and analyzed 560 phrases: 271 oral and 289 nasal phrases.

C. Instrumentation

The Velotrace, a mechanical device developed by Horiguchi and Bell-Berti (1987) for the purpose of tracking the time-varying position of the velum, was used to monitor velar kinematics. Figure 3 provides a schematic representation of the Velotrace. It consists of three major parts: a curved internal lever whose tip rests on the nasal surface of the velum, an external lever that remains in full view outside of the nose, and a push rod (carried on a support rod) that connects the internal and external levers. Because of this connection, movements of the velum that result in changes in the angle of the internal lever with respect to its fulcrum are reflected in corresponding angular movements of the external lever. The levers are connected so that when the internal lever is raised, the external lever moves toward the subject. The Velotrace has been shown to track even very rapid movements of the velum accurately (Horiguchi and Bell-Berti, 1987).

An optoelectronic system (Kay *et al.*, 1985) was used to track the movements of infrared diodes (LEDs) attached to the end of the external lever, and to the fulcrum of the Velotrace (for reference). The positions of the LEDs in the sagittal plane were tracked by a position-sensitive detector. The output was then converted into pairs of x and y coordinates for each LED. The acoustic speech signal was simultaneously recorded with the x and y coordinates onto a multichannel instrumentation recorder.

Calibration of the displacement of the external lever was accomplished by moving one diode a known distance (two centimeters) in the focal plane of the optoelectronic position sensor using a precision calibration device, and then recording its output. After digital sampling, the observed change in sampled values corresponding to this 2-cm movement was used to calibrate the recorded channels of articulator and reference signals. Note that the calibrated values of the Velotrace reflect the magnitude of movement of the external lever, rather than that of the velum itself. The external lever is about two times the length of the internal lever and so the obtained displacements are likewise larger than the actual displacements. Furthermore, it is not possible to compare the absolute magnitudes of velar gestures across subjects because, depending upon the precise positioning of the internal lever on the nasal velar surface, the magnitude of the movements of the Velotrace lever may differ.

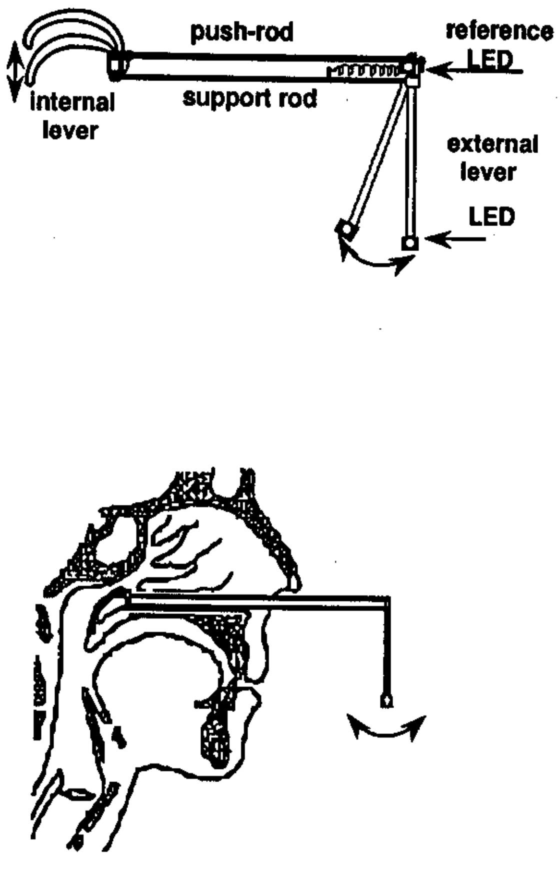

FIG. 3. Schematics of the Velotrace, above, and of the device positioned in the nasal cavity.

D. Procedure

Each subject was seated in a dental chair with a headrest adjusted to support the head in a stable and upright position. The Velotrace was positioned after the application of a topical anesthetic and decongestants to the nasal mucosa. The fulcrum of the internal lever was positioned in the nasal cavity above the end of the hard palate, with the internal lever resting on the nasal surface of the velum and the support rod on the floor of the nasal cavity. A special headband kept the Velotrace stable. A highly directional microphone was positioned in front of the subject for the purpose of obtaining a high-quality audio recording of the session, and a videotape recorder was set up to provide an audio-visual record as well. Before, during, and after the experiment, the subject was asked to produce a number of speech sounds and nonspeech postures for the purpose of checking the Velotrace signal. These included: (1) sustained /m/; (2) sustained /s/, and

(3) nasal breathing. The signals obtained during these test maneuvers were monitored on-line and recorded on FM tape for further analysis. The sustained /s/ is associated with an extreme raised position of the velum, while the sustained /m/ and nasal breathing positions are associated with extreme lowered positions. These positions provided a means for determining whether the Velotrace was tracking the full excursions of the velum in the test utterances of the experiment. We found that for one of our subjects (S3), it was not possible to identify the end of the lowering gesture for some tokens because the Velotrace signal apparently "leveled off" at the nasal breathing position. Since velar lowering offset was not one of our crucial measures and only a limited number of tokens were affected, this problem did not affect the results of the study.

E. Analysis

In the present study, we examined the vertical movements of the velum, the traditional indicator of nasal coarticulation, because they reflect changes in velar port size beyond the point at which the port is closed. The Velotrace and reference signals were digitized at 200 Hz. Once sampled, the Velotrace reference signal was smoothed (using a 25-ms smoothing window) and then subtracted from the raw Velotrace signal in order to correct for head movement. The resulting Velotrace signal (minus the reference) was smoothed with the 25-ms window as well and its velocity was obtained by taking the first derivative of the smoothed movement signal. The velocity signal was then smoothed with the standard 25-ms window. We also examined the acoustic speech signal, which was sampled at 10 000 Hz.

We identified a specific number of events in the acoustic waveform and in the movement signal (Fig. 4). For all utterances, we marked the acoustic offset of the /s/ as "s1" in "It's____" or "It's say____." We marked the acoustic onset of the /n/ as "n" in the target sequences with nasal segments and the acoustic onset of the medial /s/ as "s2" in the matched target sequences containing only oral segments. (The end of "s1" was identified as the moment at which frication noise was no longer evident in the acoustic waveform; the beginning of "n" was identified as the moment at which vowel formant energy was no longer evident in the acoustic waveform; and the beginning of "s2" was identified as the moment at which frication noise was evident.) We determined the duration of the vocalic sequence (that is, "n-s1" for the nasal utterances and "s2-s1" for the oral utterances).

We identified kinematic events using the displacement and velocity functions: Examining the displacement and velocity functions for the velum, we determined the earliest onset of velar lowering in each of the test utterances. In addition, we identified the onset of all subsequent lowering gestures [e.g., Fig. 4(b)], since many tokens showed a pattern of multistage lowering (cf. Bladon and Al-Bamerni, 1982). The method used for determining the onsets of velar lowering was derived from the commonly used approach in which movements are taken to begin and end at the time when the velocity function passes through zero. However, because it is

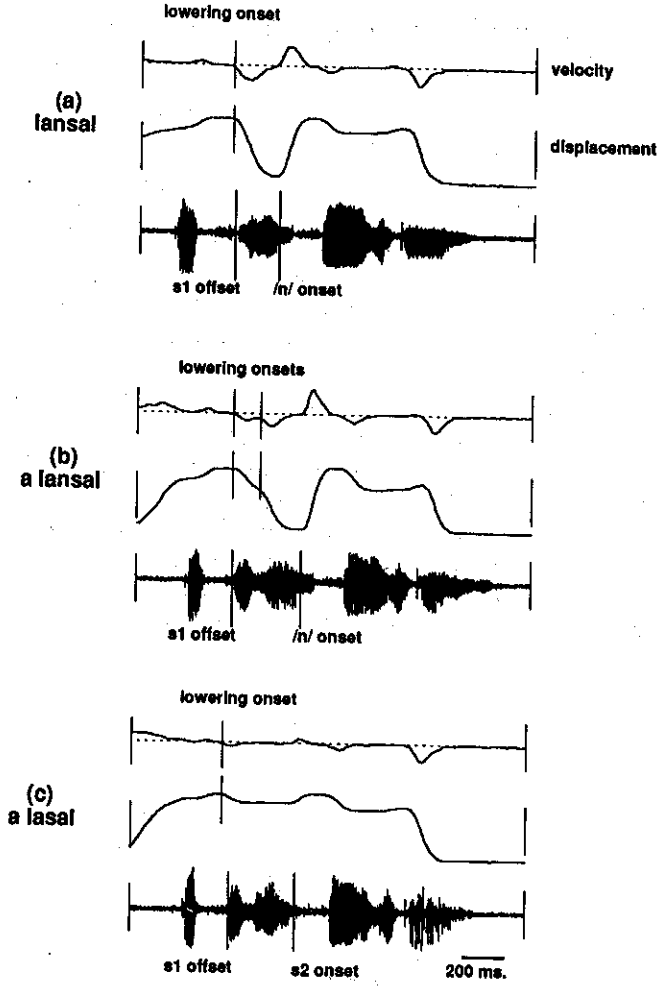

FIG. 4. Representative velocity, displacement, and acoustic signals for three tokens (two nasal and one oral), with displacement and acoustic landmarks labelled.

often the case with the velum that the velocity hovers just below or above zero (before or after passing through zero), while the associated movement appears to have ceased, strict use of zero crossings seemed inappropriate for these data. Thus, for each utterance, a negative velocity noise band below zero was used. When the velocity moved from within to below this band, a lowering movement was considered to have begun. The noise band was defined as follows. For each token, peak lowering velocity within the target item was obtained. Then, the peak values were averaged within an utterance type. For each utterance type, the noise band was defined as 10% of the average peak lowering velocity. [See Krakow (1989) for further discussion of this type of criterion.]

II. RESULTS

A. Temporal extent of velar lowering and nasal coarticulation

Traditionally, the earliest evidence of velar lowering before a $CV_n N$ sequence has been identified as the beginning of the gesture for the nasal consonant (e.g., McClean, 1973;

Moll and Daniloff, 1971), a gesture that has been shown to occur earlier in the vocalic sequence as the duration of the vocalic sequence increases. However, before linking the beginning of downward velar movements with the production of an upcoming nasal consonant, one must show that such movements do not occur in the absence of a nasal consonant. The approach we have taken here, supported by the reports of velar-position variations across oral utterances and differences among oral consonants and among vowels, was to examine minimally contrastive utterances for the presence of velar lowering for vowels preceding both nasal and oral consonants (Fig. 4).

We began our analysis by examining the nasal sequences and comparing our results with those of previous studies that had provided support for the feature spreading model. We used the procedure of those earlier studies; that is, identifying the earliest instance of velar lowering in relation to the occlusion for the nasal consonant. In agreement with those studies (e.g., McClean, 1973; Moll and Daniloff, 1971), we observed that velar lowering began earlier in advance of a nasal consonant when the preceding vocalic sequence was lengthened by the addition of segments. And, as observed in the other studies, we found a strong positive correlation ($r = 0.923$) between the duration of the vocalic sequence and the duration of the interval between the onset of velar lowering and the acoustic onset of the nasal consonant (i.e., the duration of the "anticipation") [Fig. 5(a)].

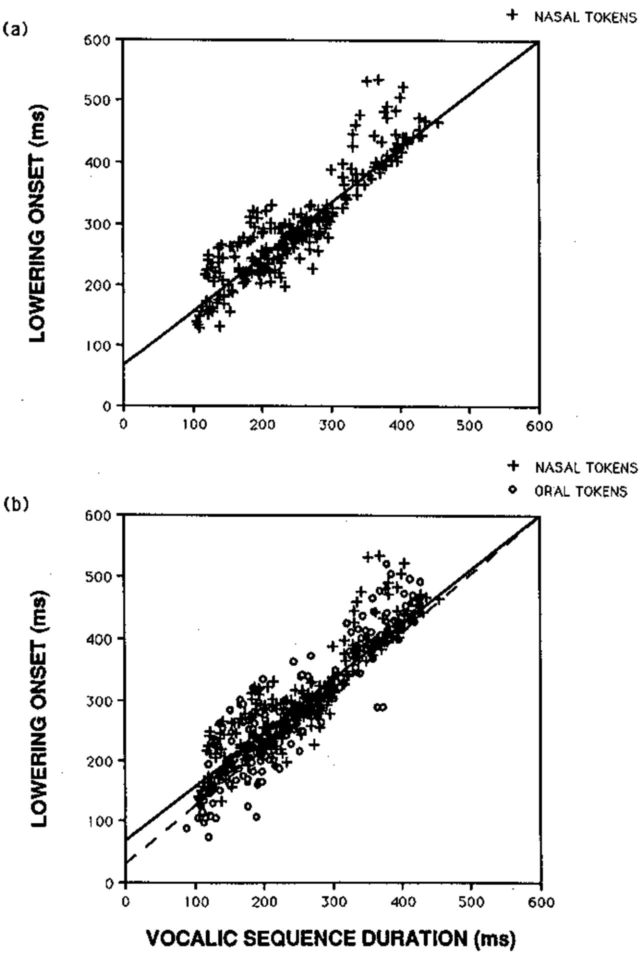

FIG. 5. (a) Scatterplots of the onset of velar lowering before /n/ versus the duration of the vocalic string, above, with data pooled across subjects ($r = 0.923$). (b) Scatterplots of the onset of velar lowering before /s/ versus the duration of the vocalic string, pooled across subjects ($r = 0.913$), superimposed on the data shown in (a).

We then turned to the oral utterances, to determine if there is velar lowering in vocalic sequences occurring in the absence of an upcoming nasal consonant, and if so, whether the temporal characteristics of the lowering gesture are similar to those of velar lowering gestures before nasal consonants. We took the equivalent measure that we used for our nasal utterances: we compared the time of the earliest velar lowering for vocalic sequences preceding the postvocalic obstruent /s/ (in the oral sequences) with the durations of those vocalic sequences. This comparison also revealed a strong positive correlation between these measures ($r = 0.913$), one that was not significantly different from the correlation for the nasal sequences [Fig. 5(a) and (b)]. The individual-subject data (Table II) also reflect the essential similarity of this measure for nasal and oral utterances.

This result reveals the error of identifying the earliest velar lowering onset in a $CV_n N$ string as anticipatory nasal coarticulation. Instead, these data suggest that at least some portion of the velar lowering movements previously attributed to feature spreading from nasal segments is, in fact, related to the articulation of the vowel string itself, since there is velar lowering in both contexts, but a nasal consonant in only one. Indeed, these results are not surprising given the cross-language evidence that the velum lowers in the transition from an oral consonant to a following vowel even in oral sequences (e.g., Bell-Berti, 1980; Bell-Berti *et al.*, 1979; Henderson, 1984; Ushijima and Sawashima, 1972).[3]

However, we also noted something about which most earlier studies have made no comment: a consistent difference in the velar lowering patterns of our sequences as a function of their length. That is, as segments were added, increasing the duration of the vocalic sequences, multistage lowering movements became increasingly evident (Fig. 6). In the longer strings, we typically saw a large velar lowering movement in close temporal proximity to the nasal consonant and a shallower lowering movement earlier in the vowel string. Note that the feature spreading approach predicts a single smooth lowering gesture; it does not predict the ap-

TABLE II. Correlations between onset of velar lowering and vocalic sequence duration.

	Nasal utterances	Oral utterances	$t =$
Subject 1	0.932	0.916	0.6347[a]
Subject 2	0.857	0.897	0.9929[a]
Subject 3 (normal rate)	0.991	0.947	4.7313[b]
Subject 3 (fast rate)	0.974	0.979	0.3446[a]

[a] $p > 0.05$.
[b] $p < 0.01$.

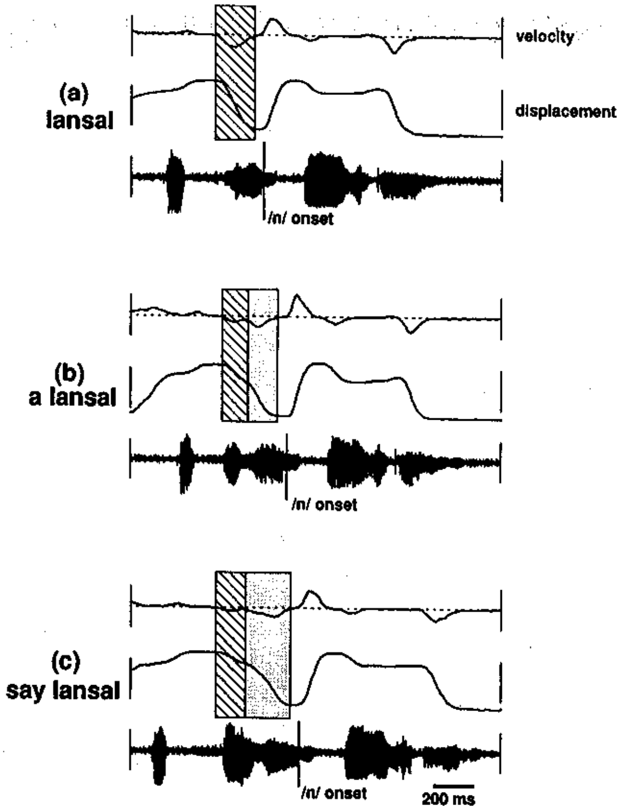

FIG. 6. Representative velocity, displacement, and acoustic signals for tokens of three nasal utterances with successively longer vocalic sequence durations achieved by the addition of segments: /lansal/, /ʌ lansal/, /seɪ lansal/. Cross-hatched rectangles indicate first (or only) component of downward movement and clear rectangles indicate second component of downward movement, identified using velocity and displacement functions (see text for complete description).

pearance of discrete vocalic and consonantal velar gestures in sequences consisting of an oral consonant, some number of vowels, and a nasal consonant. Such patterns are, however, fundamental to the coproduction approach and the particular pattern seen here is precisely what this approach predicts. That is, separately identifiable vocalic and consonantal velar gestures are expected to be evident when they have enough time in which to appear, as in sequences of longer duration.[4] In contrast, in the shorter sequences the vocalic and consonantal gestures would be predicted to overlap in such a manner that only a single lowering movement would be evident.

B. Isolating the vocalic and consonantal gestures

In order to apportion the velar lowering movement between the vocalic and consonantal articulations, we first examined the "normal" speech rate utterances of each of our subjects in which the vocalic sequences were systematically lengthened by adding segments; we then examined the "fast" utterances of our third subject to examine the effects of changing speech rate.

1. Segmental manipulation

Recall that in Fig. 6 we showed that velar lowering movements are affected by sequence lengthening due to segment addition: with increasing vocalic sequence duration, separate (vocalic and consonantal) gestures become increasingly evident as distinct movement components. Considering relative durations across the utterances of our speakers in the normal rate condition, we found that shorter utterances had only a single, smooth, lowering movement while longer utterances had two-stage lowering movements. The velocity functions reinforce this point. Note that in the longest of the three utterances shown [Fig. 6(c)], the velocity function returns to zero between the two components, whereas in the midlength utterance [Fig. 6(b)], the velocity only approaches zero between the two components, and in the shortest utterance, it reflects only a single smooth movement.

The pattern of increased separation between vocalic and consonantal gestures with increased vocalic sequence duration was evident in the data of subjects 1 and 3; for subject 2, on the other hand, we observed two-stage patterns in only a few of the longest utterances. An explanation for our cross-subject differences is that there were differences in our subjects' speech rates, with subjects 1 and 3 using a slower "normal" rate than subject 2. As shown in Fig. 2, the longest vocalic sequence durations in subject 2's utterances were substantially shorter than the corresponding utterances of subject 1 and subject 3's "normal" speech rate. In fact, the duration range of subject 2's utterances was most like that of subject 3's "fast" speech rate utterances. We argue that our results show both within- and between-subject effects of sequence duration, with subject 3 showing considerably fewer multistage velar lowering movements at the faster speech rate. The upper panels in Fig. 7 provide examples of tokens of short nasal utterances produced by each of the three subjects: Each token reveals a single large velar lowering movement. The lower panels in Fig. 7 show the movements for tokens with long vocalic sequences (lengthened by adding segments). Here, the more complex, multistage patterns are clearly evident, as are the between-subject differences. That is, two separate lowering components are readily observable in the longer tokens of subjects 1 and 3, whereas for subject 2, the component (vocalic and consonantal) gestures are not completely separate, although there is a complex lowering pattern for the longer sequence, a pattern that is quite distinct from that obtained from this subject's productions of shorter sequences. Still, it is not surprising that subject 2 showed little evidence of multistage movements since her longer sequences were similar in duration to subject 3's longer fast-rate sequences (see Fig. 2). The nature of the complexity suggests to us the presence of at least two underlying components, one (or more) vocalic, and another, consonantal.

While we have established the greater likelihood of multistage or complex gestures occurring with sequences of longer duration, we wish to strengthen our claim that the vocalic sequence is the source of the earlier stage, and that it only becomes evident when given an adequate time frame. To do this, we compared the velar displacement functions for oral-nasal minimal pairs of sufficient duration for both the vocalic and consonantal gestures to become evident. Figure 8 offers two examples of such data for longer sequences,

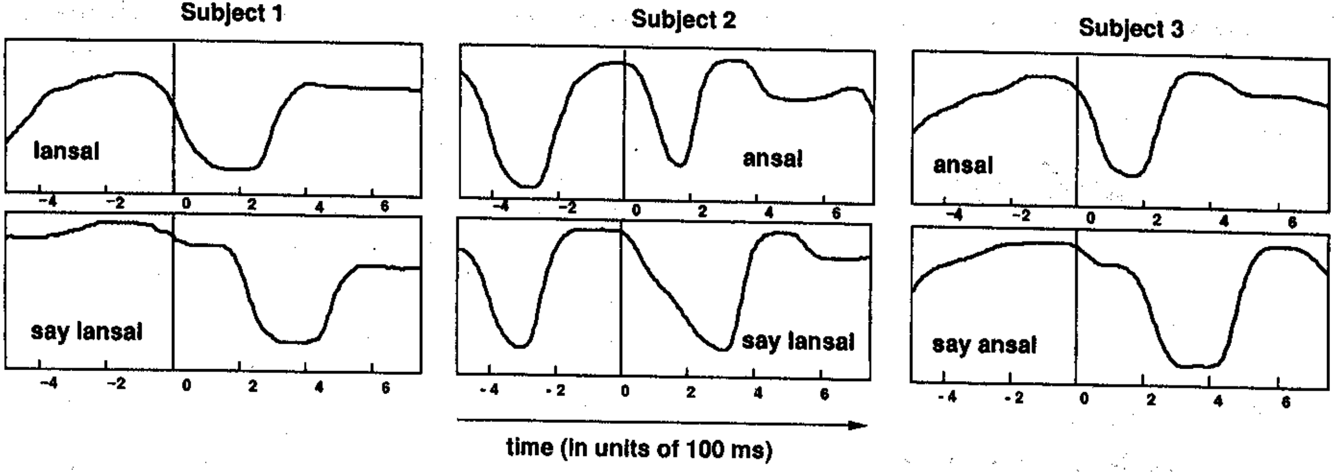

FIG. 7. Displacement functions of representative short and long nasal utterances for all three subjects, showing simple and complex movements, respectively. Displacement is represented on the ordinate, with velar lowering indicated by a downward movement. Time is represented along the abscissa, with "0" marking the end of the /s/ of the carrier phrase. (Because the data are displayed on the same time scale, we see the end velar lowering for the preceding utterance early in subject 2's data, reflecting this subject's faster speaking rate.)

in which we see two discrete movements in the nasal utterance, an early gesture of small magnitude followed by a second gesture of much greater magnitude. In the corresponding oral utterances, we would expect to see a gesture that matches that of the first stage of the complex gesture; and, indeed, we find an early gesture (of small magnitude) that is highly similar to the early movement in the longer nasal utterances.[5]

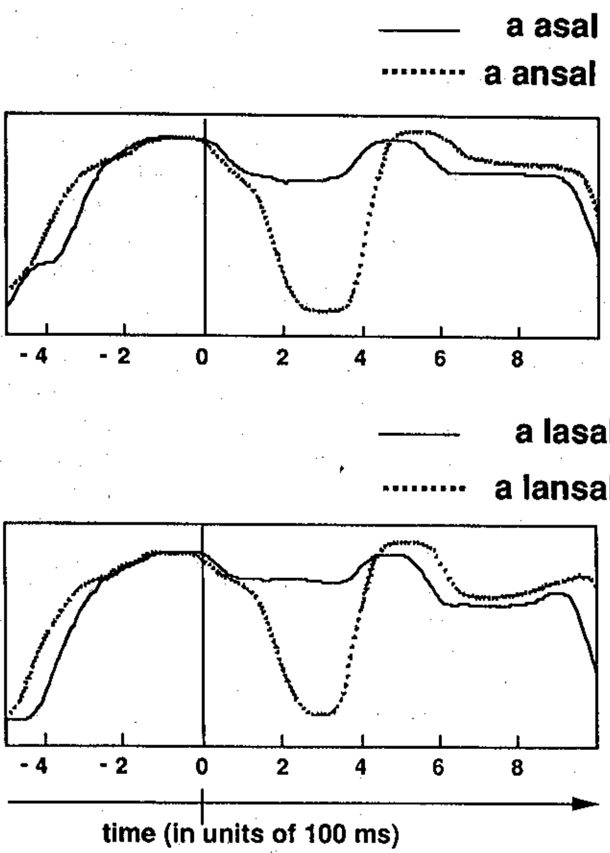

FIG. 8. Displacement functions of two representative minimally contrastive oral and nasal utterance pairs, demonstrating similar lowering onsets within each pair for subject 3. Displacement is represented on the ordinate, with velar lowering indicated by a downward movement. Time is represented along the abscissa, with "0" marking the end of the /s/ of the carrier phrase.

2. Rate manipulation

If the emergence of underlying gestures as separate on the surface is a function of reducing overlap by increasing segment duration, then the effects of speaking at a slower rate should resemble the effects of adding vocalic segments. This is not to say that the articulatory processes and strategies involved in each are the same—certainly they are not, but the two manipulations share one characteristic—that is, an increase in the duration of the vocalic string. If the predictions of the coproduction model are correct, then the result of this increase (whether due to adding vocalic segments or to speaking at a slower rate) should be a reduction in the overlap between velar lowering for the vocalic sequence and lowering for the nasal consonant. We have had an indication (by comparing the data of subject 1 and subject 3's self-selected "normal" rate with those of subject 2) that speech-rate differences can affect the amount of overlap between successive gestures. To allow us to examine the relation between speech rate and gestural organization systematically for a given subject, subject 3 was asked to produce the test utterances at a rapid speech rate in addition to the normal rate. When we explored the within-subject effects of speech rate by comparing her productions at the two different rates, we found that the slower utterances were more likely to show the multistage movements (Fig. 9). Note that both the normal and fast rate oral utterances show a vocalic velar lowering movement, but that, in these tokens, only the normal rate nasal utterance reveals separate lowering gestures for the vowels and the nasal consonant. The faster nasal utterance is simply not long enough to allow the separate components to emerge as independent. However, even in the fast rate utterances, once the vocalic portion reached some critical duration (roughly 250 ms), we did observe discrete movements

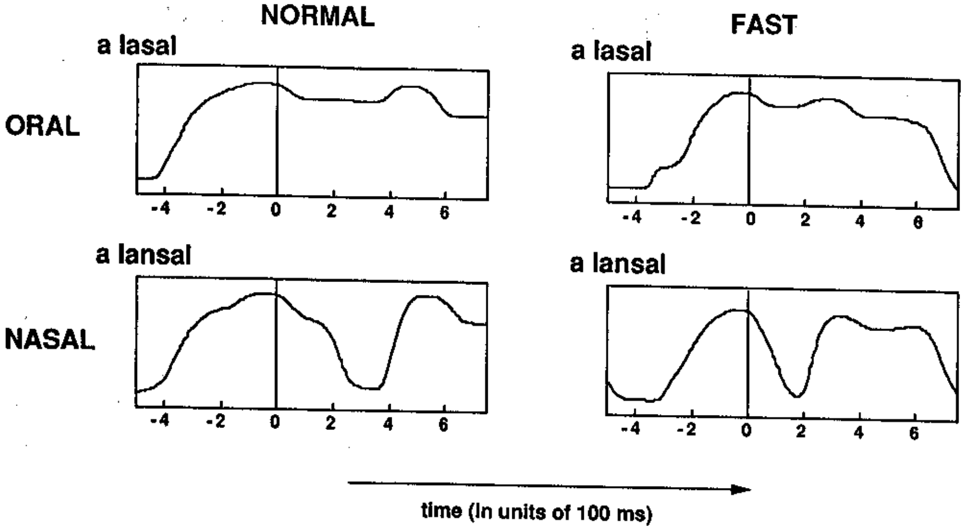

FIG. 9. Displacement functions of representative tokens of a minimally contrastive utterance pair at two speaking rates. Displacement is represented on the ordinate, with velar lowering indicated by a downward movement. Time is represented along the abscissa, with "0" marking the end of the /s/ of the carrier phrase.

for the oral and nasal portions of the utterance (Fig. 10). Thus we see the same effect of increased vowel-sequence duration, whether the increase was achieved by adding segments or by speaking at a slower rate: both manipulations resulted in an increased incidence of multistage lowering gestures. Furthermore, and as we have seen in Fig. 8 for normal speaking rate utterances, the early lowering gestures in both the normal and fast nasal utterances are paralleled in the normal and fast oral utterances (compare Figs. 9 and 10).

Figure 2 shows that the same point can be made by looking at between-subject differences in rate. To show how similar these two independent variables (rate and segment number) were with respect to this outcome, Fig. 11 shows the number of tokens with multistage lowering for subjects 1 and 3 as a function of utterance type (i.e., number of vocalic segments) and rate (only subject 3). With regard to subject 2, recall that the durations of her vocalic sequences were considerably shorter than those of subjects 1 and 3; not surprisingly, subject 2 provided scanty evidence of multistage gestures, and for this reason her data are not included in Fig. 11.[6] The incidence of multistage lowering gestures clearly increased along with the duration of the vocalic string. Least likely to show multistage movements were the faster (rate), shorter (number of segments) sequences and, most likely, were the slower, longer sequences.

To quantify the effect of vocalic sequence duration on the number of stages observed in the lowering gestures, we

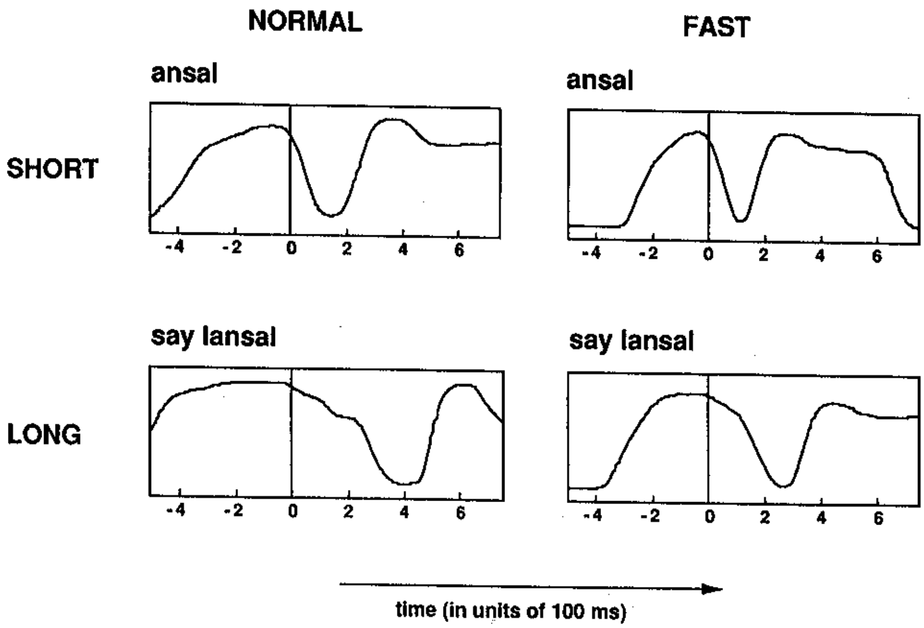

FIG. 10. Displacement functions of representative tokens of short and long utterances at two speaking rates for subject 3. Displacement is represented on the ordinate, with velar lowering indicated by a downward movement. Time is represented along the abscissa, with "0" marking the end of the /s/ of the carrier phrase.

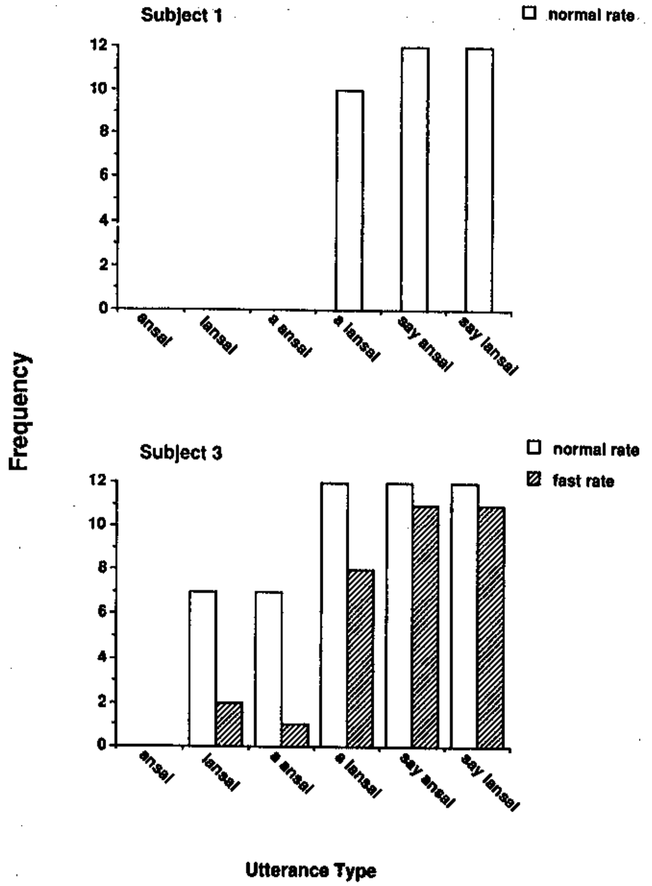

FIG. 11. Histogram of frequency of multistage gestures in nasal utterances of subject 1 and subject 3.

pooled the data for our subjects and divided the nasal utterances into two groups, those having a single stage and those having more than one stage. We also divided the utterances into two groups on the basis of vocalic sequence duration, a group of "short" utterances (those falling below the median duration of 241 ms) and a group of "long" utterances (those falling above the median duration). The resulting frequency distribution is given in Table III.

III. DISCUSSION

In feature spreading models, anticipatory velar lowering is predicted to extend back to the first vocalic segment in a string preceding a nasal consonant. That is, the motor plan for a nasal consonant is presumed to change with changes in the context in which it is embedded. Moreover, the effect of the nasal consonant is theoretically unbounded. In coproduction models, anticipatory velar lowering is simply viewed as an effect of the temporal overlap of the intrinsic velar gestures for adjacent vowels and consonants. That is, each segment, including the vowels, is specified as having an associated velar gesture, and the influence of a segment on its neighbors is presumed to be limited in duration.

TABLE III. Observed frequency of single and multiple gestures as a function of vocalic sequence duration (short or long).

	Short	Long	Totals
One	118 (74.68%)	40 (25.32%)	158
Multi	27 (20.61%)	104 (79.39%)	131
Totals	145	144	289

This experiment is designed to compare the predictions of feature spreading and coproduction models systematically for sequences consisting of an oral consonant followed by some number of vowels and a nasal consonant. Using minimally contrastive sequences of the form CV_nC and CV_nN has allowed us to clarify the role of articulatory movements intrinsic to a particular segment and the coarticulatory influence of phonetic context in producing observed velar displacement patterns, and has provided strong support for the coproduction model of Bell-Berti and Harris (1981).

The results indicate that much of what has been attributed to the coarticulatory influence of a nasal consonant is actually velar lowering that would be observed in any CV_n sequence, independent of the presence of an upcoming nasal consonant. It should have come as no surprise to researchers that the velum lowers in the transition between an oral consonant and an oral vowel, since it has long been known that vowels are produced with their own intrinsic velar positions and that these positions are lower than those observed for oral consonants. With a short vocalic sequence between an oral consonant and a nasal consonant, it is difficult to separate the vocalic and nasal consonantal influences, because they appear as one continuous lowering movement. However, when a vocalic string is lengthened, either by segment addition or by speaking at a slower rate, separate vocalic and nasal consonantal movements emerge. What is more, the relatively shallow early part of the lowering movement observed in the CV_nN sequences matches that observed in minimally contrastive CV_nC sequences that end with an oral, rather than a nasal, consonant. Those ending with a nasal consonant, of course, also contain a sharp, extensive velar lowering gesture; however, this latter movement occurs at a relatively stable time in close temporal proximity to the acoustic onset of the nasal murmur and, by inference, the achievement of oral closure.

Previous research, carried out without minimally contrastive oral sequences, has led to the earliest onset of velar lowering being identified as the onset of nasal coarticulation. This approach has its origin in the phonological notion of segment underspecification and the specific hypothesis that lacking an oral/nasal contrast, English vowels are unspecified for velar position and, thus, strongly affected by their consonantal context. Underspecification is a crucial notion in phonology and also, it has been suggested, in some phonetic domains [see Keating (1988) for a recent discussion of this issue]. However, our results indicate that it is an inappropriate and misleading notion when applied to the organization of articulatory gestures, including velar movement [see Boyce *et al.* (1991) for a discussion of the same issue concerning lip rounding]. Moreover, even though phonological descriptions need only specify a binary distinction between "oral" and "nasal," the articulatory organization requires n-ary (i.e., multivalued) distinctions of velar height that also clearly have phonological implications. For example, the relation between vowel height and velar height is relevant to an understanding of why, in the languages of the world, distinctive vowel nasalization is more commonly found in low than in high or mid vowels (see Beddor, 1983).

Returning for a moment to the frequent observation that velar height for vowels is lower than that for oral consonants, we believe that this relation has explanatory power for reconciling the data of Bladon and Al-Bamerni (1982) with our own. They claimed to have found a combination of feature spreading and coproduction strategies that, we believe, is compatible with the predictions of the coproduction model alone if intrinsic velar positions for vowels and the effects of variable overlap of vocalic and consonantal gestures are included in the analysis. Their primary argument for concluding that speakers combine both feature spreading and coproduction strategies was their observation of movements in which there was two-stage lowering: a first stage that began near the release of the initial (oral) consonant in a CV_nN string, and a second stage that was closely timed to the onset of the nasal consonant occlusion. Since the first stage appeared earlier in advance of the nasal consonant occlusion as the vowel string increased in duration, they considered this as evidence for feature spreading. The second stage was considered as evidence for time locking (i.e., coproduction), because of its constant and relatively close temporal relation to the nasal consonant occlusion. However, this is precisely the pattern one would expect if the early, stage 1, movement reflects lowering for the vowel and is unrelated to the upcoming nasal consonant, and the stage 2 movement reflects lowering for the nasal consonant. This is the pattern that we observed when we lengthened the vocalic string sufficiently for the observation of discrete vocalic and nasal consonantal velar gestures.

Bladon and Al-Bamerni were also puzzled by the alternation between such two-stage movements and single-stage lowering that appeared smooth and continuous and whose onset occurred earlier in relation to the nasal consonant as the vocalic string was lengthened. We also found this alternation; in our data, its source can be clearly seen to reside in changes in the duration of the vocalic string that led to more or less overlap between vocalic and nasal consonant gestures and, thus, more or less observed separation of the two lowering components. Unfortunately, it is not possible to reconcile our interpretation of this pattern with that of Bladon and Al-Bamerni because their report is not sufficiently detailed to make it possible for us to determine the durations of their CV_nN strings. We would obviously expect the patterns to be correlated with durational differences.

In our discussions, we have assumed that separate vocalic and nasal consonant gestures are present in the shorter or faster utterances, but that they are coincident, or, at least, overlap substantially, and are, therefore, not independently observable (Fig. 10). We want to make it clear that we also believe that the vocalic lowering we have observed includes separate components for the individual vowels in the sequence, although the present experiment was not designed to separate those from one another. Nonetheless some of our slower and longer utterances showed multiple shallow gestures in the vocalic sequence (i.e., Fig. 9—long normal rate utterance). The velum has often been referred to as a "slow" articulator, and Bell-Berti (1980) reported data suggesting that, all else being equal, movements for a segment begin about 250 ms before the oral articulation. Our present data support the view that there is a limited and stable timing relation between the onsets of velar movements and the corresponding oral gestures. It is, perhaps, not surprising, then, that one does not find evidence of each of the vocalic gestures in the displacement pattern, since few of them achieved individual durations of such length. Furthermore, to make such effects evident one must use segments of clearly differing intrinsic velar positions; this, in turn, requires further study, to determine, for example, the intrinsic positions for the liquid and glide segments that were used in creating our long vocalic sequences. Since we found that the number of gestures in a displacement trajectory varied with speaking rate and the length of the sequence (in segments), we suggest that the timing of velar gestures is stable across speaking rates. We therefore suppose that if very slow speech were studied, one would find more separate gestures in the displacement pattern.

IV. CONCLUSION

To reconcile the disagreements about the adequacy of the two types of coarticulation models, we conducted a carefully controlled study of velar movements in CV_nC and CV_nN sequences. The results indicate that the interpretation of the earliest onset of velar lowering in CV_nN strings as coarticulation of the nasal consonant is unfounded. That is, similar lowering occurs in strictly oral sequences. Instead, these data show that there are intrinsic velar positions for the vocalic sequence and for the upcoming nasal consonant, each of temporally limited extent. Observed patterns of coarticulation, then, are simply the result of the temporal overlap of the nasal consonant gesture with the gestures for the vocalic sequence. The portion of the vocalic sequence that is overlapped by the nasal consonant gesture increases as the duration of the vocalic sequence decreases. This study strongly supports the coproduction model and shows how certain misinterpretations of data have led to the conflicting conclusions.

ACKNOWLEDGMENTS

This work was supported by NIH Grant DC-00121 to the Haskins Laboratories. We wish to thank Katherine Harris, Ignatius Mattingly, Leigh Lisker, Mary Boyle, and Carol Fowler for their helpful comments on an earlier version of this manuscript.

[1] C is an oral consonant, V_n is any number of vowels, and N is a nasal consonant.

[2] The languages studied include English, Hindi, German, and Japanese.

[3] Such patterns are also seen in CV sequences containing high and mid vowels (Bell-Berti, 1980; Bell-Berti *et al.*, 1979; Henderson, 1984).

[4] These results also reflect the report of Bladon and Al-Bamerni, although they do not report consistency of this pattern, nor do they comment on the relation between vocalic sequence duration and number of stages.

[5] We expect the slope of the early stage observed in longer nasal utterances to become more like that of the oral utterances as vocalic sequence duration increases, pushing the onset of the nasal consonant gesture further from the onset of the (first) vocalic gesture, which would reduce the gestural overlap earlier in the vocalic sequence.

[6] The few tokens in which multi-stage gestures are found are, however, tokens of "long" utterances.

Beddor, P. S. (**1983**). "Phonological and phonetic effects if nasalization on vowel height," Ph.D. thesis, University of Minnesota, Minneapolis, Minnesota (reproduced by the Indiana University Linguistics Club).

Bell-Berti, F. (**1980**). "A spatial-temporal model of velopharyngeal function," in *Speech and Language: Advances in Basic Research Practice* (Vol. IV), edited by N. J. Lass (Academic, New York).

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

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

Bell-Berti, F., and Harris, K. S. (**1982**). "Temporal patterns of coarticulation: Lip rounding," J. Acoust. Soc. Am. **71**, 449–454.

Benguerel, A.-P., and Cowan, H. A. (**1974**). "Coarticulation of upper lip protrusion in French," Phonetica **30**, 41–55.

Benguerel, A.-P., Hirose, H., Sawashima, M., and Ushijima, T. (**1977**). "Velar coarticulation in French: A fiberoptic study," J. Phon. **5**, 149–158.

Bladon, R. A. W., and Al-Bamerni, A. (**1982**). "One-stage and two-stage temporal patterns of velar coarticulation," J. Acoust. Soc. Am. Suppl. 1, **72**, S104.

Boyce, S. E. (**1988**). "The influence of phonological structure on articulatory organization in Turkish and English: Vowel harmony and coarticulation," unpublished Ph.D. thesis, Yale University, New Haven, Connecticut.

Boyce, S. E., Krakow, R. A., and Bell-Berti, F. (**1991**). "Phonological underspecification and speech motor organization," Phonology (in press).

Boyce, S. E., Krakow, R. A., Bell-Berti, F., and Gelfer, C. E. (**1990**). "Converging sources of evidence for dissecting articulatory movements into core gestures," J. Phon. **18**, 173–188.

Browman, C. P., and Goldstein, L. P. (**1986**). "Towards an articulatory phonology," Phonology Yearbook **3**, 219–252.

Daniloff, R., and Moll, K. (**1968**). "Coarticulation of lip rounding," J. Speech Hear. Res. **11**, 707–721.

Fowler, C. A. (**1980**). "Coarticulation and extrinsic theories of timing control," J. Phon. **8**, 113–133.

Fritzell, B. (**1969**). "The velopharyngeal muscles in speech: An electromyographic and cineradiographic study," Acta Otolaryngologica, Suppl. 250.

Gelfer, C. E., Bell-Berti, F., and Harris, K. S. (**1989**). "Determining the extent of coarticulation: Effects of experimental design," J. Acoust. Soc. Am. **86**, 2443–2445.

Harris, K. S. (**1970**). "Physiological aspects of articulatory behavior," Haskins Labs. Status Rep. on Speech Res. **23**, 49–67.

Henderson, J. B. (**1984**). "Velopharyngeal function in oral and nasal vowels: A cross-language study," unpublished Ph.D. thesis, University of Connecticut, Storrs, Connecticut.

Henke, W. (**1966**). "Dynamic articulatory model of speech production using computer simulation," unpublished Ph.D. thesis, MIT, Cambridge, Massachusetts.

Horiguchi, S., and Bell-Berti, F. (**1987**). "The Velotrace: A device for monitoring velar position," Cleft Palate J. **24**, 104–111.

Joos, M. (**1948**). "Acoustic phonetics," Language **24**, 1–136.

Kay, B. A., Munhall, K. G., Vatikiotis-Bateson, E., and Kelso, J. A. S. (**1985**). "Processing movement data at Haskins: Sampling, filtering, and differentiation," Haskins Labs. Status Rep. on Speech Res. **81**, 291–303.

Keating, P. A. (**1988**). "Underspecification in phonetics," Phonology **5**, 3–29.

Kent, R. D., Carney, P. J., and Severeid, L. R. (**1974**). "Velar movement and timing: Evaluation of a model of binary control," J. Speech and Hear. Res. **17**, 470–488.

Kent, R. D., and Minifie, F. D. (**1977**). "Coarticulation in recent speech production models," J. Phon. **5**, 115–133.

Kozhevnikov, V., and Chistovich, L. A. (**1965**). "Speech: Articulation and Perception," Joint Publications Res. Serv., Washington, DC.

Krakow, R. A. (**1989**). "The articulatory organization of syllables: A kinematic analysis of labial and velar gestures," unpublished Ph.D. thesis, Yale University, New Haven, Connecticut.

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

Liberman, A. M., Cooper, F. S., Harris, K. S., and MacNeilage, P. F. (**1962**). "A motor theory of speech perception," Proc. Speech Commun. Seminar, Stockholm: Royal Inst. Technol.

McClean, M. (**1973**). "Forward coarticulation of velar movements at marked junctural boundaries," J. Speech Hear. Res. **16**, 286–296.

Moll, K. L. (**1962**). "Velopharyngeal closure of vowels," J. Speech Hear. Res. **5**, 30–77.

Moll, K. L., and Daniloff, R. G. (**1971**). "Investigation of the timing of velar movements during speech," J. Acoust. Soc. Am. **50**, 678–684.

Munhall, K. G., and Lofquist, A. (**1991**). "Gestural aggregation in speech: Laryngeal gestures," J. Phon. (in press).

Ohala, J. J. (**1971**). "Monitoring soft palate movements in speech," Project on Linguistic Analysis Reports, Phonology Laboratory, Department of Linguistics, University of California Berkeley, J01-J015.

Öhman, S. E. G. (**1966**). "Coarticulation in VCV utterances: Spectrographic measurements," J. Acoust. Soc. Am. **39**, 151–168.

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

Perkell, J. S. (**1986**). "Coarticulation strategies: Preliminary implications of a detailed analysis of lower lip protrusion movements," Speech Commun. **5**, 47–68.

Perkell, J. S., and Chiang, C.-M. (**1986**). "Preliminary support for a "hybrid model" of anticipatory coarticulation," Proc. 12th Int. Congr. Acoust. July 1986, A3-6.

Saltzman, E. L., and Munhall, K. G. (**1989**). "A dynamical approach to gestural patterning in speech production," Ecol. Psychol. **1**, 333–382.

Schourup, A. (**1973**). "A cross-language study of vowel nasalization," Ohio State Univ. Working Papers in Linguistics **15**, 190–221.

Ushijima, T., and Hirose, H. (**1974**). "Electromyographic study of the velum during speech," J. Phon. **2**, 315–326.

Ushijima, T., and Sawashima, M. (**1972**). "Fiberscopic examination of velar movements during speech," Ann. Bull. Res. Inst. Logoped. Phoniatr., 25–38.