/003

Influences on articulatory timing in consonant sequences

Dani Byrd*

UCLA Department of Linguistics

This paper presents an experimental study of articulatory timing in English consonant sequences using electropalatography (EPG). Effects of consonantal place, manner, and syllable structure on reduction and temporal overlap are considered. The results evidence reduction in coda position and show that stops are more subject to coda reduction than are fricatives. Coronal consonants are shown to be more overlapped by a following velar stop than a velar stop is by a following coronal. Stops are more overlapped by a following consonant than are fricatives. Finally, an onset cluster is shown to be less overlapped and less variable in its timing than coda clusters and heterosyllabic sequences. These findings support the claim that inter-gestural coordination is variable and affected by linguistic factors—both gestural and prosodic.

© 1996 Academic Press Limited

1. Introduction

Currently, one of the most significant challenges in the study of speech production is to gain a theoretical understanding of how speakers coordinate articulatory movements. The goal of this effort is to uncover principles of coordination rather than simply patterns of coordination. Many factors are known to affect articulatory timing but rarely have these effects been studied with respect to consonant sequences or with articulatory data collected from multiple speakers. Such a study is reported here. Consonant sequences are of special interest in creating models of speech production, as often many demands are concurrently placed on an individual articulatory structure, the tongue. The tongue must execute these demands in a short period of time, and the consonants are not discretely articulated. Although consonant sequences are of special importance in understanding articulatory organization, little articulatory data has been published on such sequences. Most articulatory research on coproduction has considered single intervocalic consonants. Hypotheses regarding the degree of coarticulation in consonant sequences, as compared to adjacent consonants and vowels, must be evaluated empirically. Consonant cluster timing is likely to be variable and subject to myriad influences interacting in complex ways.

*Now at Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, U.S.A., where correspondence should be addressed.

Some work can be found on the nature of coarticulation in consonant sequences, but the coverage of this work is not complete. In English, phonetic observation has revealed that the closure for the first consonant in a cluster generally is not released until after the closure for the second is formed (e.g., Kenyon, 1951; Jones, 1956; Abercrombie, 1967; Catford, 1977; MacKay, 1978; Hardcastle & Roach, 1979; Ladefoged, 1993; see Henderson & Repp, 1982 on acoustic consequences). Barry (1985), Nolan (1992), Browman & Goldstein (1990b), and others have shown that an initial consonant in a cluster that may sound as if it has been assimilated in place of articulation, or deleted altogether, is often in fact still articulated at its original point of articulation. The articulations of consonants in sequences overlap (e.g., Recasens, Fontdevila & Pallarès, 1993) and may also reduce or be incomplete (Barry, 1985, 1991; Kerswill, 1985; Nolan, 1992). Byrd (1992) used articulatory synthesis to show that a completely articulated alveolar stop is not perceived by listeners if it is substantially overlapped with a velar stop. This and other work indicate that purely perceptual or acoustic approaches to understanding these speech events may not always be revealing when consonant clusters are involved. Detailed information is needed on articulation to determine how such sequences are coordinated.

The experiment described here examines the spatial and temporal production of English alveolar and velar stops and alveolar fricatives in sequence. Sequences of the type [C#C], [#CC], and [CC#] are considered, where # is a word boundary. Electropalatography (EPG) is used to quantify how the coproduction of consonant sequences varies as a function of the consonants' place and manner, and the placement of syllable boundaries. The term coproduction is used here and throughout simply to refer to the temporal co-occurrence or overlap in the articulation of two (or more) gestures.

The discussion of the results adopts the framework of Articulatory Phonology as a starting point because it offers an explicit approach to characterizing speech timing through the postulation of phasing relations (Browman & Goldstein, 1986, 1988, 1989, 1990a,b, 1992a, and elsewhere). That is, it postulates the existence of abstract gestural primitives that are coordinated with one another such that a phase angle in one gesture is specified as synchronous with a phase angle in another gesture. The determination of inter-gestural timing in terms of phasing relations is discussed further in Sections 4 and 5, and in Byrd (in press). Examination of some of the factors relevant to intergestural timing is the focus of this experiment.

Additionally, consonant reduction is investigated concomitantly with the timing issues. The term reduction is used in the discussion below to refer to changes in magnitude of lingua-palatal contact *or* in duration of contact for a particular consonant as a function of syllable position. That is, a consonant may reduce spatially or temporally. The experiment examines how the syllabic position of a consonant, i.e., word-final (coda) *vs.* word-initial (onset), influences degree and duration of lingua-palatal contact. For example, a tendency to decrease the gestural magnitude of codas is often assumed to be the motivation for lenition processes. The following hypothesis regarding coda reduction is proposed:

(H1) Consonants have less lingua-palatal contact in coda position than in onset position.

Coda reduction has been shown for labial and coronal stops (see e.g., Fromkin, 1965; Krakow, 1989; Browman & Goldstein, 1992b, 1995). We will consider

the coronal fricative and the velar stop to determine if this is generalizable to other places and manners, as is suggested by results reported by Browman & Goldstein (1995). *Variability* in magnitude and duration is examined to determine whether it is greater in contexts having a decreased mean displacement or duration. For example, assuming the validity of the preceding hypothesis (H1), is it true that:

(H2) Lingua-palatal contact is more variable for consonants in coda position than in onset position.

As noted above, the relative timing of sequences of two consonants is of primary interest here. It bears mentioning that degree of articulatory overlap and variability in that overlap are conceptually and empirically independent issues, just as spatial displacement and variability are. Two gestures could, for example, potentially be relatively un-overlapped and very stable, or relatively overlapped and very stable. For this reason, degree of coproduction *and* timing stability are both considered here.

In the first portion of this experiment, sequences of two consonants in different order across a word boundary (e.g., "bag dab" & "bad gab") are compared. Hardcastle & Roach (1979) found that the time between the initiation of closure for an adjacent [t] and [k] was shorter for a [tk] cluster than for a [kt] cluster in a front vowel context. This suggests the formulation of a hypothesis regarding the effect of place on articulatory overlap:

(H3) A tongue tip gesture is more temporally overlapped with a following tongue body gesture than a tongue body gesture with a following tongue tip gesture.

This hypothesis is relevant to phonological issues because of the exceptional behavior of coronals in assimilation processes and the often-made, controversial assumption that this asymmetry is due to the fact that coronals are underspecified for Place features. H3, if supported, suggests an articulatory basis for this behavior (see also Byrd, 1992; Barry, 1992; and Browman & Goldstein, 1992a).

Because there is reason to believe that place effects on timing exist, and because of the potential interaction of overlap and aerodynamics in the case of continuant consonants, articulatory manner effects are also investigated. The degree of overlap permitted in a particular sequence may be a function of the constriction degree of the consonants involved, with less overlap expected in sequences including a fricative, as a coarticulated closure would impede the airflow necessary for frication. The following hypothesis is tested with the [C#C] sequences:

(H4) A closure gesture is more temporally overlapped by a following consonant articulation than is a fricative gesture.

Understanding the interaction of prosodic structure such as syllable boundaries and articulation is a crucial question for both segmental and dynamic phonological theories. The final section of this experiment considers the effects of the placement of syllable (word) boundaries on gestural magnitude and inter-gestural timing by comparing onset clusters, coda clusters, and heterosyllabic sequences. These data are relevant to ongoing research efforts to determine the effects of prosody on articulation (e.g., Hardcastle, 1985; Beckman, Edwards & Fletcher, 1992; Pierrehumbert & Talkin, 1992). The following two hypotheses are tested to determine how syllable affiliation affects relative timing in clusters:
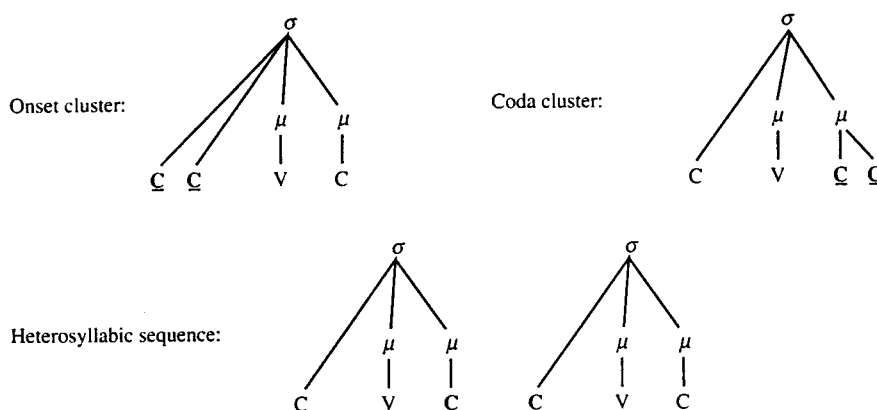
Figure 1. Possible representations of onset, coda, and heterosyllabic consonant sequences.

(H5) Temporal coproduction in consonant sequences is greater if the consonants are tautosyllabic, less if they are heterosyllabic.

(H6) Temporal coproduction in consonant sequences is less variable if the consonants are tautosyllabic than if they are heterosyllabic.

The latter has been found to be the case by Browman & Goldstein (1988) and others.

Work in prosodic phonology (McCarthy & Prince, 1986; Hayes, 1989; Zec, 1989) has led to the proposal of certain constituent structures for coda, onset, and heterosyllabic sequences. As shown in Fig. 1, in these works consonants in a coda cluster (upper right of figure) are superordinated by a single mora (μ), in turn dominated by the syllable node (σ). Sequences of onset consonants (upper left of figure), however, are represented as being individually directly dominated by the syllable node. Sequences of consonants spanning a word boundary (lower portion of figure) would be dominated jointly only by phrasal constituents.

Given certain assumptions about the relationship between prosodic constituency and phonetic timing, these representations would be predicted to accord with particular articulatory patterns. That is, it seems not unreasonable to expect those consonants exhaustively forming a constituent, like the coda cluster, to exhibit more coproduction than those not comprising a constituent, like the onset and heterosyllabic sequences. This suggests a simple extension of hypothesis five (H5: overlap is greater within syllables than across syllables) to the moraic constituent. Given these structures and the above assumption, the extended hypothesis, H5', is that the coda cluster, forming a moraic constituent, is more overlapped than the other sequence types. Additionally, it might be assumed that those consonants forming a moraic constituent are more cohesive, i.e., more stable in their timing, than the other sequences. Likewise, this assumption and the structures in Fig. 1 suggest an extension of H6. Just as timing is hypothesized to be more stable within a syllable (H6), it is by extension hypothesized, H6', to be more stable within a coda cluster. This experiment considers the degree of compatibility, given these assumptions regarding timing and constituent structure, between the articulatory data and the prosodic representations shown in Fig. 1.

2. Method

2.1. *Stimuli*

Two-member consonant sequences are evaluated here. (For a consideration of longer sequences, see Byrd, 1995b). Throughout this study, word boundaries are used as a general diagnostic for syllable boundaries. (Hardcastle & Roach (1979) found no effect of a word boundary (-VC#CV'-) as opposed to a coda + onset boundary (-VCCV-) on their timing measures.) The experimental sequences are shown with the frame sentences in Table I (underlining added). Control utterances of homorganic clusters, discussed later, are also shown in Table 1.

The vowels bordering the consonants were [æ]'s; these in turn were bordered by bilabial stops to minimize any lingual coarticulation with the target clusters. The use of the adjacent low front vowel was intended to create a somewhat front velar stop constriction that would be most observable on the EPG pseudopalate and, at the same time, to minimize vocalic lingua-palatal contact. As /æ/ cannot occur at the end of a word, the initial sequence /#sk/ was preceded by [ə]. Also, as /gs#/ is not possible in English, the voicing of one of the consonants had to be changed to create an eligible coda cluster. The material was randomized in 10 blocks; thus 10 repetitions were recorded. Tokens from blocks two through eight are included in the data analysis.

2.2. *Speakers*

Five speakers were recorded and paid at a standard compensation rate. Speakers will be referred to as Speakers A, B, K, M, and S. These include two men (M & S) and three women who have grown up and been educated in Southern/Central California. All speakers speak a dialect typical of this area; none reported any speech or hearing pathology.

2.3. *Data collection*

Electropalatography (EPG) uses an artificial palate of thin acrylic embedded with electrodes. The Kay Elemetrics Palatometer system, employed here, uses a

TABLE I. Consonant sequences and frame sentences (underlining added)

Consonants C1 C2	Heterosyllabic sequences C#C	Onset cluster #CC	Coda clusters CC#
d g	*Type bad gab again.*		
g d	*Type bag dab again.*		*Type bagged amp again.*
s g	*Type bass gab again.*		
g/k s	*Type bag sab again.*		*Say backs Abigail.*[1]
s k	*Type bass cap again.*	*Type a scab again.*	*Type mask amp again.*
control sequences:			
d d	*Type bad dab again.*		
s s	*Type bass sa again.*		
g g	*Type bag gab again.*		

Where # stands for a word (syllable) boundary.

[1] The tokens used for [ks#] were taken from the recordings made for another experiment which used a slightly different frame sentence.

pseudopalate of thin acrylic, extending around the teeth, with 96 electrodes. The Palatometer scans the palate at a 100 Hz sampling rate with a scan time of 1.7 ms to acquire all 96 values in a sample. The palate electrodes were calibrated with the Palatometer software.

Before the experiment, the speakers wore their artificial palates for an hour of normal activity to accommodate. Subjects were seated in the UCLA Phonetics Laboratory near the Palatometer computer module, facing away from the monitor. One practice page of material was read by the speaker before the recording started. Nonsense words (e.g., *sab*) were pointed out to the speaker at this time. Speakers were cued for each sentence by the word "Go" from the experimenter. Each sentence was cued individually, and there was a pause after each one. Subjects were instructed to maintain a constant, fluent reading rate with no unusual stresses, and deviations from this were pointed out by the investigator during the training block. If a speaker paused, hesitated, or otherwise had a false start, he or she was re-cued for a repetition of that sentence. All material was recorded in one two-hour session. Speakers were required to take a break half-way through the session.

2.4. *Data analysis*

Articulatory regions on the pseudopalate were determined empirically for each speaker. (For a discussion of the rationale for speaker-specific regions and for EPG indices, see Byrd, Flemming, Mueller, & Tan, 1995). This determination was based on the homorganic control utterances having no lingual coarticulation with an adjacent consonant, 10 tokens of [d#d], [s#s], and [g#g] sequences for each speaker.[2] All tokens of these homorganic sequences were used to establish regions on their pseudopalates which will be referred to as "front" and "back". Crucially, for each subject, no electrodes that were contacted at the frame of minimum contact (fewest number of electrodes contacted) for the vowel, [æ], were included in the consonantal regions. All electrodes contacted after the vowel minimum, i.e., during the consonant formation, until and including the frame of maximum contact during the consonant, were assigned to the relevant region—front for [s] & [d], back for [g]. Any electrodes which were designated in this way as members of both regions were also excluded. These cases were generally few and always adjacent to the electrodes excluded as vocalic contact locations. All other (i.e., uncontacted) electrodes were included in the region to which they were physically closest, as determined by measurements made with a flexible ruler on the acrylic pseudopalates. (Diagrams of the resulting regions for each speaker can be found in Byrd & Tan, this issue.) This protocol ensures that the moment of initial regional contact observed for the consonant sequence will in fact be the concomitant of the formation of C1, as opposed to the vowel articulation or C2. That is, measurements made from examining front and back regional contact over time are conservative in identifying temporal edges of consonantal contact, but there is a high degree of confidence that the contact measured is actually attributable to the consonant in that region.

[2] One potential limitation of EPG is that electrode coverage is limited mostly to the hard palate area making it possible that contact on the pseudopalate under-represents the full area of velar closure, specifically contact occurring well onto the soft palate. However, this problem is not severe for front velars. An examination of ten repetitions of the control utterances for velars ([g#g]), showed that every token for every subject had a complete seal across the back of the palate for the velar closure. Some tokens showed up to five electrodes contacted along the mid-sagittal plane.
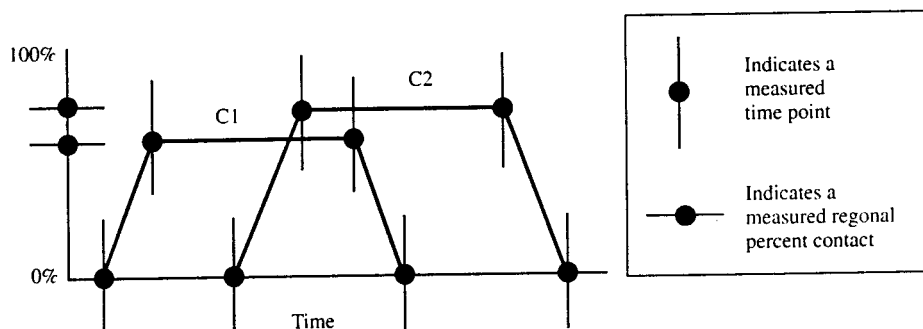
Figure 2. Schema of a [CC] contact profile showing four measured timepoints during lingua-palatal contact for each consonant and maximum regional lingual-palatal contact for each consonant.

For each consonant sequence token, these pseudopalate regions were used to determine the percentage of electrodes in the region having lingua-palatal contact at each sample. Temporal displays, called "contact profiles", that show the total number or percent electrodes contacted at each frame over time can be used as the basis for quantitative evaluation of EPG data (e.g., Barry, 1991). This study uses a percent display that shows the percent electrodes contacted in the front and back regions in each frame rounded to the nearest integer. These displays were used to determine quantitative measures of articulatory timing and magnitude. The contact profiles for both regions were examined to determine four timepoints during each period of consonant activity: first frame with any contact in the region, first frame at maximum contact, last frame at maximum contact, and last frame with any contact in the region. (Note that between the first and last frame of maximum contact, any dips away from maximum contact, although infrequent, were allowed.) The maximum percent regional contact was also recorded. Fig. 2 shows a schematic contact profile for a CC sequence. In the schema, first, regional contact (front for [d] or [s], back for [k] or [g]) for C1 increases, and then, during the period of maximum contact for C1, contact for C2 initiates. The schema in Fig. 2 shows the four time-points measured for each consonant and the maximum regional contact percentage also measured for each consonant. These measures were used, as described below, to calculate indices (similar to those used by Barry, 1991) reflecting the latency and overlap between the two consonants, and the duration of contact and maximum contact for each individual consonant.

The specific indices reflecting *duration* are:[3]

- Region duration (FRONT DURATION and BACK DURATION)—the duration of lingua-palatal contact in a region in seconds; (*time of final contact—time of initial contact*)$_{region}$;
- Sequence duration (SEQUENCE DURATION)—total duration of linguapalatal contact for a sequence in seconds; (*time of final contact − time of initial contact*)$_{sequence}$.

The index reflecting a consonant's amount of *contact* is:[4]

- Maximum contact (FRONT MAXIMUM & BACK MAXIMUM)—the maximum percent contact in a region.

[3] These and additional EPG indices are also discussed in Byrd *et al.* (1995).
[4] For other indices reflecting the shape of a consonant's EPG contact profile see Byrd *et al.* (1995).

Maximum lingua-palatal contact is indicative of a consonant's degree of lingual displacement. Reduced consonants will have less contact.

The indices reflecting the temporal *overlap* in contact for the tongue tip and tongue body consonants are:

- Sequence overlap (SEQUENCE OVERLAP(%))—the percentage of the total sequence duration during which contact occurred in both regions;
- C1 overlap (C1 OVERLAP (%))—the percentage of C1 duration during which contact for C2 also occurred;
- C2 overlap (C2 OVERLAP (%))—the percentage of C2 duration during which contact for C1 also occurred.

SEQUENCE OVERLAP indicates the degree of coproduction occurring in the sequence. C1 OVERLAP and C2 OVERLAP indicate how overlapped a particular consonant is by another constriction.

Latency refers to the time interval between two specific articulatory events, such as the onset of contact for each consonant. The indices reflecting the temporal *latency* between tongue tip and tongue body contact are:

- Time between onsets (ΔONSETS)—time between initial contact in one region and initial contact in the other region in seconds;
- Time between peaks (ΔPEAKS)—time between maximum contact in one region and peak contact in the other in seconds, where the time of maximum contact is calculated as the temporal center of any plateau of maximum contact;
- Time between C1 release and C2 onset (ΔC1 RELEASE TO C2 ONSET)—time between the first frame after the final frame of maximum contact for C1 and the initial contact for C2 in seconds;
- C2 onset relative to C1 (C2 ONSET RELATIVE TO C1 (%))—the percentage of the way through C1 at which the initial contact for C2 occurs;
- C2 peak relative to C1 (C2 PEAK RELATIVE TO C1 (%))—the percentage of the way through C1 at which the maximum contact for C2 occurs, where the time of maximum contact is calculated as any temporal center of the plateau of maximum contact.

It has been proposed by Browman & Goldstein (1990b) that consonants in sequence are phased so that the second consonant begins at the offset (290°) of the preceding one. A consistent value of 0 for the index ΔC1 RELEASE TO C2 ONSET would correspond to such a phasing relationship. (Of course our observed onset will be later than the gestures' actual onsets, as is true to a certain extent with all kinematic measurements, movement tracking or otherwise.) The variables of C2 ONSET RELATIVE TO C1[5] and C2 PEAK RELATIVE TO C1 are also measures of latency, these taking into account differences in C1 duration.

Data analysis focuses on tendencies apparent across subjects; however, significant individual differences are also reported. The spatial and temporal indices outlined above are used in a repeated-measures General Linear Model (GLM) Analysis of

---

[5] Note that C2 ONSET RELATIVE TO C1 and C1 OVERLAP are not redundant because it is possible for initial C2 contact to precede C1 contact.

Variance (ANOVA) model. Except where otherwise noted, this model uses pooled data with [Speaker] added as a random independent variable. This computational method uses the [Speaker × Variable] interaction as the error term in the test for [Variable] as described by Winer (1971) to provide control over individuals between experimental units (see Choi, 1992). (Thus, the degrees of freedom reported for the error term are those of the [Speaker × Variable] interaction.) The SuperAnova package (Abacus Concepts, 1989) is used to perform the statistical tests. If there is no significant main effect but four of the five speakers show the same direction of effect, the main effect is recalculated excluding the exceptional speaker. This is reported when significant, and the excluded speaker is noted. This procedure was adopted to ensure that robust behaviors were not ignored simply because a single speaker's means deviate hugely from the others. Probabilities less than or equal to 0.05 are considered significant; probabilities less than or equal to 0.08 are noted as non-significant trends (*ns trend*). Because variability is also of interest, the Levene statistic for testing equal variability will also be employed (Levene, 1960 cited in Dixon, 1988). This statistic uses the absolute values of the deviations from the group means as data. The deviations were calculated here separately for each subject using his or her mean for the variables: both front and back regions' MAXIMUM and DURATION, and each sequences' C2 ONSET RELATIVE TO C1 and ΔC1 RELEASE TO C2 ONSET. The Levene *F* statistic was then computed as a one-way ANOVA *F* using the computational method for repeated-measures outlined above. It is recognized that one of the assumptions underlying the ANOVA is that variance within each of the treatment groups is homogeneous. However, *F*-tests are in fact quite robust to departures from homogeneity of variance (Winer, 1971). In the rare instances yielding a significant main effect on variability as determined by the Levene statistic and in means as determined by ANOVA, the means for the treatment levels are given for comparison purposes.

3. Results

3.1. *Place, manner, and syllable position effects*

First, effects of sequence order on overlap and reduction in the heterosyllabic sequences [d#g], [g#d], [s#g], and [g#s] are reported. Because of the inappropriateness of post-hoc tests for repeated measures analyses, the pairs of sequences— [d#g] *vs.* [g#d] and [s#g] *vs.* [g#s]—are analyzed separately.

First, consider the stop–stop sequences. The contact profiles for these sequences are shown in Fig. 3A–J and are presented in pairs for each speaker in order to facilitate comparisons of the sequences. The left panels show front-back sequences, and the right panels show back-front sequences. If there were *no* overlap, we would expect in the front-back sequences to see the profiles with circles (front region contact) rise and fall, and only then the profiles with squares (back region contact) rise and fall. With no overlap, the back-front sequences' profiles would have squares before circles. Rather, notice that all profiles exhibit overlap, with much variability in the onset of C2 with respect to C1.

Generally, we observe that coda consonant articulations are reduced and variable relative to onsets and that the onset of C2 (relative to the onset of C1) is later for [g#d] than for [d#g]. In fact, for most of the speakers, the consonants of [d#g] are
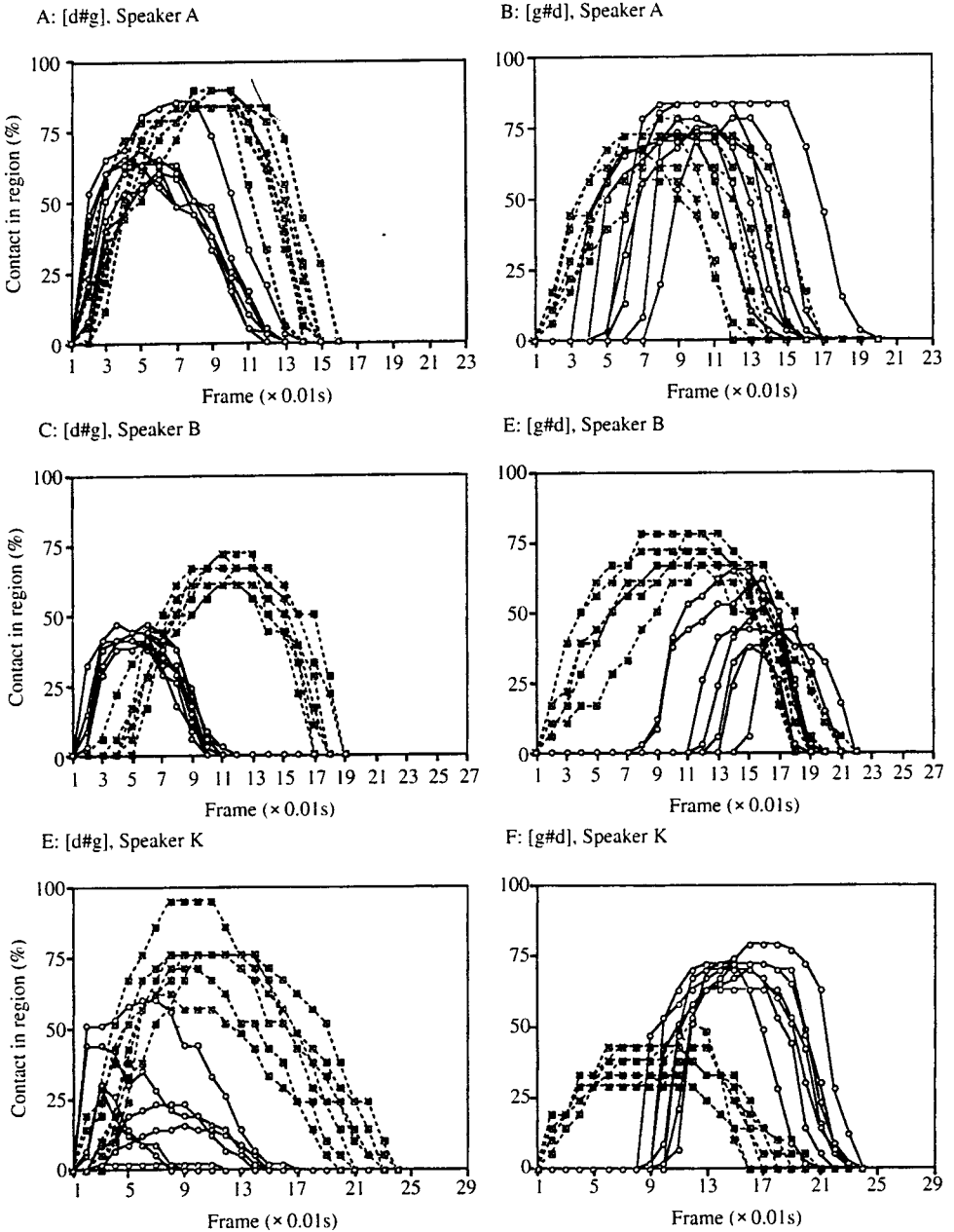
A: [d#g], Speaker A

B: [g#d], Speaker A

C: [d#g], Speaker B

E: [g#d], Speaker B

E: [d#g], Speaker K

F: [g#d], Speaker K

Figure 3. Contact profiles for [d#g] (left) and [g#d] (right) for five speakers.
—○—, front region; ···□···, back region.

nearly completely overlapped, with contact for [g] often starting synchronously with that for [d]. Table II summarizes the results of the ANOVA comparing [d#g] and [g#d].

These results demonstrate that both [d] and [g] have less displacement in coda than in onset position. The maximum contact differs in the two sequences, having
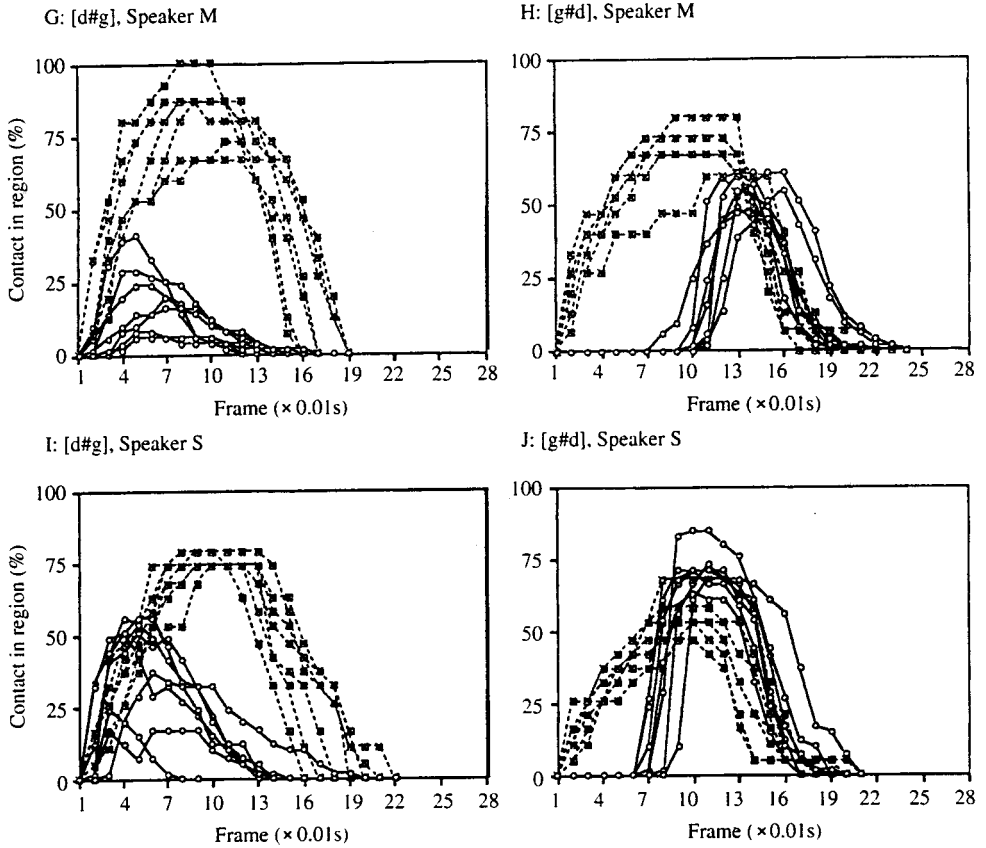
G: [d#g], Speaker M

H: [g#d], Speaker M

I: [d#g], Speaker S

J: [g#d], Speaker S

Figure 3. *Continued.*

TABLE II. A summary of the significant effects in the statistical analysis of [d#g] and [g#d]. (*ns trend* indicates a non-significant trend of $p \leq 0.08$)

| Dependent variable	Confidence level $F(1, 4)$, $F(1, 3)$ when 1 Sp. excluded	Description
Individual Consonants		
FRONT MAXIMUM	$F = 13.156$; $p = 0.0222$	[g#d] > [d#g]
BACK MAXIMUM	$F = 25.252$; $p = 0.0152$	[d#g] > [g#d], Speaker B excluded
Timing		
SEQUENCE DURATION	$F = 9.184$; $p = 0.0563$; *ns trend*	[g#d] > [d#g], Speaker S excluded
SEQUENCE OVERLAP (%)	$F = 6.865$; $p = 0.0790$; *ns trend*	[d#g] > [g#d], Speaker B excluded
C1 OVERLAP (%)	$F = 58.425$; $p = 0.0016$	[d#g] > [g#d]
ΔONSETS	$F = 50.625$; $p = 0.0021$	[g#d] > [d#g]
C2 ONSET RELATIVE TO C1 (%)	$F = 35.631$; $p = 0.0040$	[g#d] > [d#g]
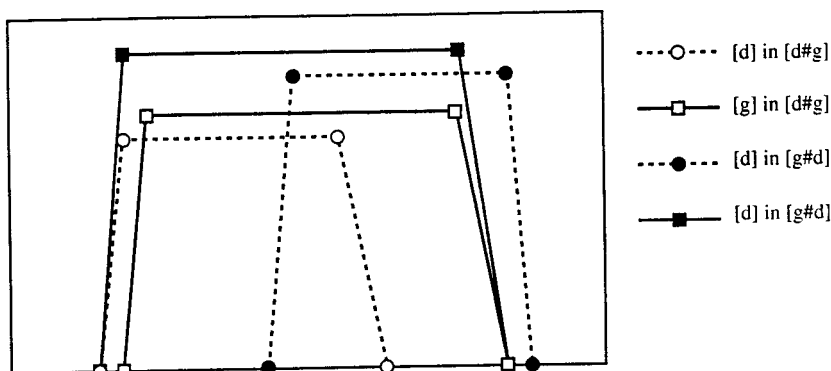
Figure 4. Schematic of the timing relationship for [d#g] and [g#d]; to scale for the pooled mean values of ΔONSETS, FRONT DURATION, and BACK DURATION only.

the following means: [d] in coda 40% and in onset 65%; [g] in coda 61% and in onset 78%.

Next, consider the coproduction between the two consonants. As is evident from even a casual inspection of the contact profiles, [d#g] sequences are substantially more overlapped than [g#d] sequences. This timing difference is responsible for the effects detailed in the lower half of Table II. The percentage of the sequence during which contact in both regions occurs (SEQUENCE OVERLAP) and the overlap of C1 by C2 (C1 OVERLAP) are both much greater for [d#g] than for [g#d]. The mean SEQUENCE OVERLAP is 59% for [d#g] and 46% for [g#d]. The mean C1 OVERLAP is 87% for [d#g] and 53% for [g#d]. The difference in coproduction is also evidenced by the greater latency between onsets for [g#d] and the fact that C2 starts much later relative to C1 for [g#d]. The mean value for the time between onsets is 0.01 s (1 frame) for [d#g] and 0.07 s (7 frames) for [g#d]. The measure of relative latency, C2 ONSET RELATIVE TO C1, has means of 8% for [d#g] and 46% for [g#d]. Finally, total sequence duration is shorter in [d#g] (0.15 s versus 0.17 s). This is not surprising in light of the extensive overlap in this sequence. Consider the schematic of the timing relationship in the stop–stop sequences in Fig. 4 which is to scale for the *pooled* mean values of ΔONSETS, FRONT DURATION, and BACK DURATION *only*.

While the duration of contact for each consonant is not significantly different as a function of sequence, the shortness of the [d] causes it to be proportionally more overlapped than [g] regardless of its syllable position. The absolute latencies also differ substantially. Virtually all of the contact for [d] occurs during that for [g], except for a slight latency between the onsets of C1 and C2 contact which preserves canonical order. While [g#d] has a greater latency between onsets, it has a slight offset between the moments of *final* contact.

The other pair of sequences relevant in considering place and manner effects are [s#g] and [g#s]. Speaker M's contact profiles for these sequences are shown as examples in Fig. 5. (In the interest of space, only a subset of the profiles are presented for many of the sequences in the remainder of Section 3. Those shown are representative of the patterns exhibited across subjects.) Statistical analysis comparing [s#g] and [g#s] is summarized in Table III.

Asymmetries between the front-back and back-front order are in the same direction as found for the stop–stop sequences. The [g] contact is greater in onset
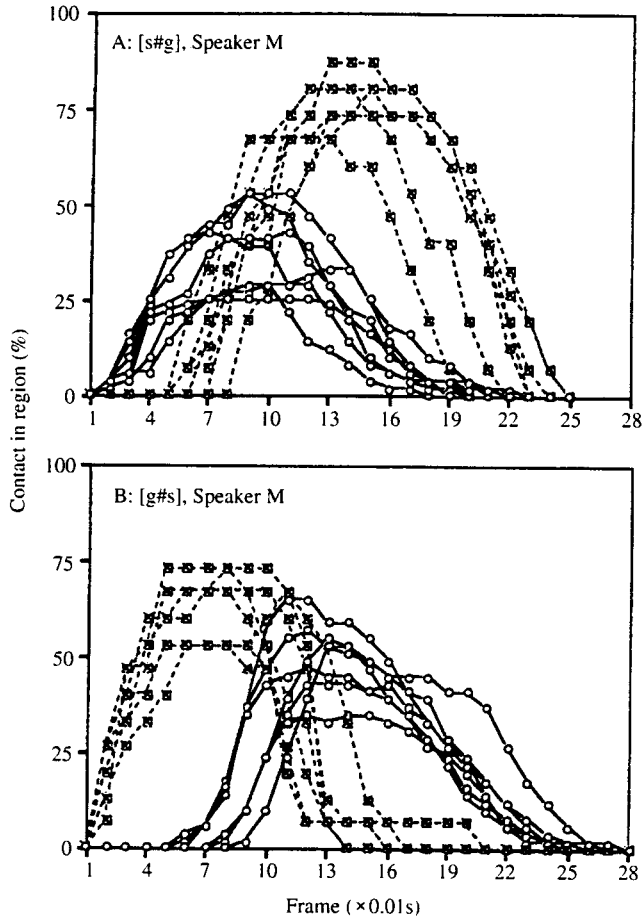
Figure 5. Contact profiles for Speaker M for (A) [s#g] and (B) [g#s].

position, having a maximum of 64% in [s#g] and 50% in [g#s]. There is a tendency for all speakers except Speaker S to have both shorter [s] and [g] contact in coda position. This parallels results in Byrd (1995a). Three speakers do show somewhat lower maximum contact for [s] in coda position as compared to onset position. (For [b#s] and [s#b] sequences, Byrd (1995a) finds [s] contact duration to be shorter in coda but finds no difference in maximum contact.) Like the stop–stop sequences, these sequences are significantly more overlapped when the front consonant precedes the back than the reverse. Consistent differences in absolute latency of C2 are not so readily apparent. Recall that the phasing relationship proposed by Browman & Goldstein (1990b) is the synchronization of the onset of C2 with the release of C1. This measure shows no regular difference between [d#g] and [g#d], although there is a significant interaction of main effect and speaker with three speakers having greater intervals for [g#d]. This measure does distinguish [s#g] and [g#s] which are not distinguished by ΔONSETS as [d#g] and [g#d] are. The *pooled* means for the timing measures are given in Table IV; the stop–stop means are included for comparison.

Next, recall that the effect of manner on overlap—that is, whether a stop is more

TABLE III. A summary of the significant effects in the statistical analysis of [s#g] and [g#s]. (*ns trends* indicates a non-significant trend of $p \leq 0.08$)

Dependent variable	Confidence level $F(1, 4)$; $F(1, 3)$ when 1 Sp. excluded	Description
Individual Consonants		
BACK MAXIMUM	$F = 16.18$; $p = 0.0276$	[s#g] > [g#s], Speaker B excluded
BACK DURATION	$F = 8.142$; $p = 0.0649$	[s#g] > [g#s], Speaker S excluded
Timing		
SEQUENCE DURATION	$F = 13.073$; $p = 0.0364$	[g#s] > [s#g], Speaker K excluded
SEQUENCE OVERLAP (%)	$F = 14.601$; $p = 0.0188$	[s#g] > [g#s]
C1 OVERLAP (%)	$F = 9.650$; $p = 0.0530$; *ns trend*	[s#g] > [g#s], Speaker A excluded
C2 OVERLAP (%)	$F = 67.755$; $p = 0.0012$	[s#g] > [g#s]
ΔC1 RELEASE TO C2 ONSET	$F = 14.678$; $p = 0.0313$	[s#g] > [g#s], Speaker M excluded
C2 ONSET RELATIVE TO C1 (%)	$F = 9.650$; $p = 0.0530$; *ns trend*	[s#g] > [g#s], Speaker A excluded
C2 PEAK RELATIVE TO C1 (%)	$F = 16.573$; $p = 0.0152$	[g#s] > [s#g]

overlapped by a following consonant than is a fricative—is of interest. An ANOVA including the sequences [d#g] and [s#g] determines that the stop–stop sequence exhibits significantly more overlap of C1 by C2 ($F(1, 4) = 12.327$; $p = 0.0247$). The onset of C2 relative to C1 is earlier for [d#g] ($F(1, 4) = 14.141$; $p = 0.0198$) and the time between onsets is less ($F = 34.778$; $p = 0.0041$). There is also a trend for the interval from the release of C1 to the onset of C2 to be greater for [s#g] than [d#g] ($F(1, 3) = 10.095$; $p = 0.0502$, Speaker B excluded). These measures all indicate greater coproduction in the stop–stop sequence. When the effect of the manner of C2 for the sequences in the reverse order—[g#d] and [g#s]—is considered, somewhat parallel results are found. SEQUENCE OVERLAP ($F(1, 4) = 13.938$, $p = 0.0202$), and C2 OVERLAP ($F(1, 4) = 110.708$, $p = 0.0005$) both indicate greater

TABLE IV. A summary of means for [g#d], [d#g], [g#s], and [s#g], pooled across speakers; * indicates a significant difference as shown in Table II or Table III, *tr.* indicates a trend ($p \leq 0.08$)

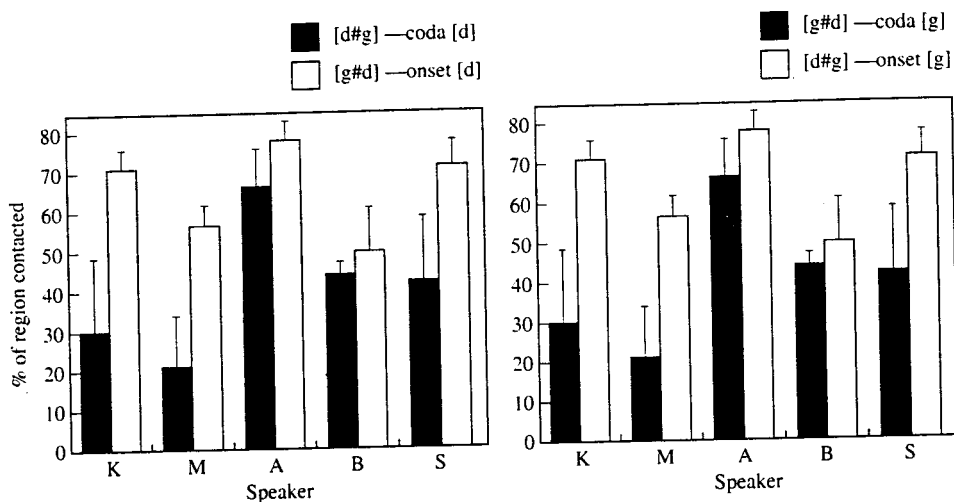|	[s#g]		[g#s]	[d#g]		[g#d]
SEQUENCE DURATION	0.20 s	*	0.22 s	0.15 s	*tr.*	0.17 s
SEQUENCE OVERLAP	44%	*	27%	59%	*tr.*	46%
C1 OVERLAP	55%	*tr.*	50%	87%	*	53%
C2 OVERLAP	66%	*	37%	62%		80%
ΔONSETS	0.07 s		0.06 s	0.01 s	*	0.07 s
ΔC1 RELEASE TO C2 ONSET	−0.01 s	*	−0.03 s	−0.04		−0.03
C2 ONSET RELATIVE TO C1	45%	*tr.*	50%	8%	*	46%
C2 PEAK RELATIVE TO C1	82%	*	99%	86%		77%

Figure 6. Maximum contact and standard deviation in the stop–stop sequences for [d] (left panel) and [g] (right panel).

overlap in the stop–stop sequence. There is also a trend for [g#s] to have longer absolute latencies of C2 (ΔPEAKS & ΔC1 RELEASE TO C2 ONSET) than [g#d], (Speaker B being exceptional). The general conclusion is that the sequences including a fricative are less overlapped than those having only stops.

The formulation of experimental hypotheses outlined an interest in the role of variability in reduction. Specifically, given the spatial and temporal reduction seen in coda position, codas are hypothesized to be more variable than onsets. Recall that in the stop–stop sequences both consonants reduce in MAXIMUM contact in coda position. Fig. 6 shows maximum contact and standard deviation for each speaker's consonants in the stop–stop sequences.

A consequence of visually examining the contact profiles is that one is left with the distinct impression that certain dimensions of particular sequences are much more variable than others. The Levene statistic, described in Section 2.4, is used to compare variability; significant effects are reported. The Levene analysis confirms that for four speakers [d] is more variable in MAXIMUM contact in coda position than in onset position ($F(1, 3) = 13.099$; $p = 0.0363$, Speaker B excluded). In the analysis of the fricative [s], there are also parallels between reduction and increased variability, in this case a lack of both. Recall that for the two sequences [g#s] and [s#g], there was no significant difference in [s] MAXIMUM or DURATION. The Levene analysis also finds no significant differences in variability in these dimensions. However, coda [g] did reduce in back MAXIMUM contact, but neither pair of sequences has greater overall variability in this measure for coda [g]. Finally, C2 ONSET RELATIVE TO C1 is also more variable for [g#s] than [s#g], with Speaker K being the only exceptional speaker. This will be discussed further below.

3.2. *Sequence type effects*

Recall that overlap and variability as a function of prosodic structure is of interest in the latter portion of this experiment which includes onset clusters, coda clusters,

and heterosyllabic sequences. Articulatory timing, variability, and reduction are considered in the sequences: [s#k], [#sk], [sk#], [g#d], [gd#], [g#s], and [ks#]. Each subgroup, that is, fricative–stop, stop–stop, and stop–fricative, will be examined separately. First, however, timing is considered for all seven sequences together.

An ANOVA was conducted for which the seven sequences were coded as heterosyllabic ([s#k], [g#d], [g#s]), onset ([#sk]), or coda ([sk#], [gd#], [ks#]). Overlap between the consonants was found to differ depending on sequence type: the first consonant in an onset cluster is less overlapped by a following consonant than it is in a coda cluster or heterosyllabic sequence ($F(2, 6) = 8.128$; $p = 0.0196$, speaker B excluded). ΔONSETS and ONSET OF C2 RELATIVE TO C1 are distinguished similarly with the onset cluster having a longer latency than the coda or heterosyllabic sequences ($F(2, 8) = 10.26$; $p = 0.0062$ and $F(2, 6) = 8.802$; $p = 0.0164$, Speaker B excluded, respectively). The coda and heterosyllabic sequences consistently differ in SEQUENCE DURATION ($F(2, 8) = 12.492$; $p = 0.0035$) which increases from coda to heterosyllabic to onset. They also differ consistently in ΔPEAKS ($F(2, 8) = 34.643$; $p = 0.0001$) and ΔC1 RELEASE TO C2 ONSET ($F(2, 8) = 50.115$; $p = 0.0001$) but in opposite directions. While the onset cluster has longer latencies in both cases, the coda cluster has longer ΔC1 RELEASE TO C2 ONSET than the heterosyllabic sequence but shorter ΔPEAKS. Finally, the *second* consonant in a coda cluster is less overlapped than in an onset or heterosyllabic sequence ($F(2, 6) = 6.311$; $p = 0.0334$, Speaker M excluded). This may be due to some word-final lengthening which extends C2, thereby decreasing the proportion of it overlapped by C1.

To examine differences in contact profiles for individual consonants due to syllabic affiliation and to further explore the timing differences, each of the subgroups is considered separately, i.e., [s#k], [#sk], [sk#] and [g#d], [gd#] and [g#s], [ks#].

Representative contact profiles for the [s-k] sequences are shown in Fig. 7A–C for Speaker K. Both spatial and temporal organization are important considerations. Significant main effects from the ANOVA for these three sequences are reported in Table V.

If it is simply the case, based on the results of Section 3.10 for example, that articulations are of smaller magnitude when in coda position, then reduction of C1 would be expected to pattern alike in [sk#] and [s#k] since it is in coda position in both cases. If the relevant factor for reduction is whether the consonant is word-final, then, for example, [k] might reduce in [sk#] but not in [s#k] or [#sk]. The differences in front region contact duration show that [s] is shortest as a single consonant in coda position and longest when part of an onset cluster. Being the first consonant in a coda cluster results in a duration intermediate between that of the single coda consonant and the first consonant in an onset cluster. The back consonant is shortest as the second member of the onset cluster, but shows no significant changes in degree of lingua-palatal contact. For four speakers, differences among the sequences in BACK MAXIMUM were negligible; one speaker (A) had a somewhat lower value for the heterosyllabic sequence.

If the degree of cluster coproduction is determined by whether the consonants occur in the same syllable, [#sk] and [sk#] would be expected to group together in coproduction. We do see a substantial effect of sequence type on the timing measures. Onset clusters are significantly less overlapped than coda and heterosyllabic sequences and have longer absolute and relative latencies. The onset cluster is also longer in total duration than heterosyllabic sequences and coda clusters. This is
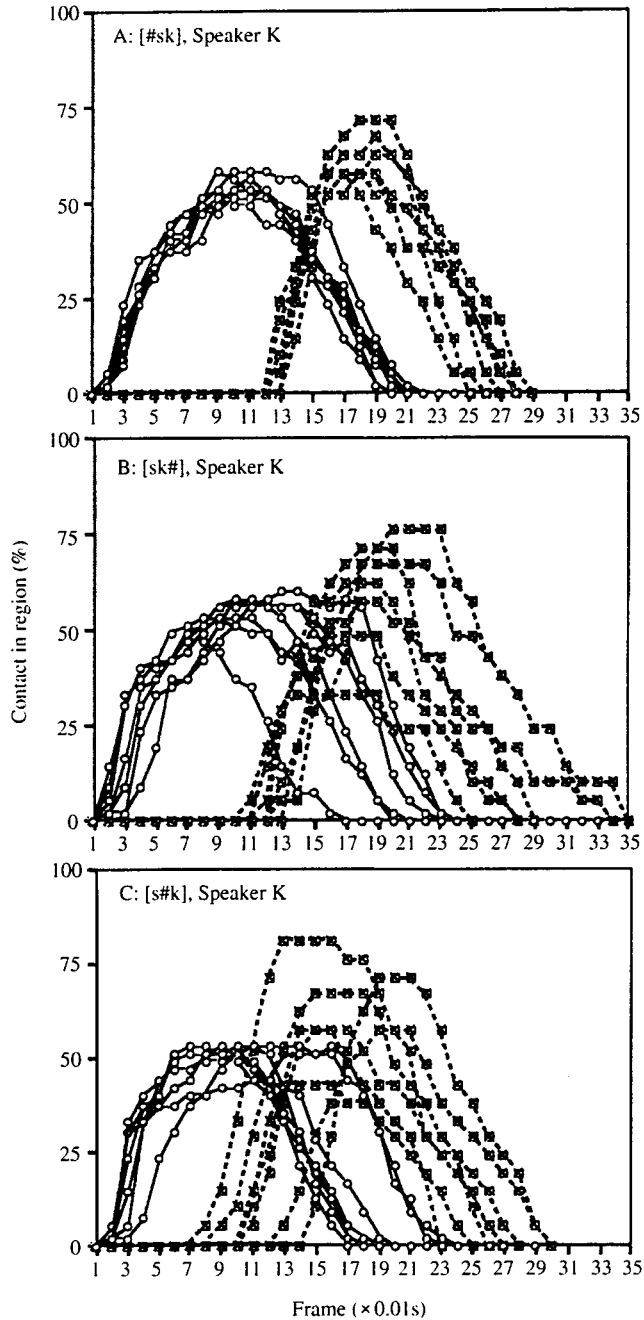
Figure 7. Contact profiles for Speaker K for (A) [#sk], (B) [sk#], and (C) [s#k].

in accordance with the decreased overlap in these clusters. However, there are no consistent differences in timing between the coda cluster and heterosyllabic sequence. The pooled means for the timing measures are given in Table VI. Although not consistent across speakers, there is a general tendency for the coda

TABLE V. A summary of the significant effects in the statistical analysis of [s#k], [#sk], [sk#]

Dependent variable	Confidence level $F(2, 8)$, $F(2, 6)$ when 1 Sp. excluded	Description
Individual Consonants		
FRONT DURATION	$F = 4.565$; $p = 0.0476$	[#sk] > [sk#] > [s#k]
BACK DURATION	$F = 6.491$; $p = 0.0316$	[#sk] < [sk#] & [s#k], Speaker M excluded
Timing		
SEQUENCE DURATION	$F = 9.023$; $p = 0.0155$	[#sk] > [sk#] & [s#k], Speaker K excluded
SEQUENCE OVERLAP (%)	$F = 4.747$; $p = 0.0437$	[#sk] < [sk#] & [s#k]
C1 OVERLAP (%)	$F = 11.717$; $p = 0.0042$	[#sk] < [sk#] & [s#k]
ΔONSETS	$F = 11.928$; $p = 0.0040$	[#sk] > [sk#] & [s#k]
ΔPEAKS	$F = 4.872$; $p = 0.0413$	[#sk] > [sk#] & [s#k]
ΔC1 RELEASE TO C2 ONSET	$F = 6.263$; $p = 0.0231$	[#sk] > [sk#] & [s#k]
C2 ONSET RELATIVE TO C1 (%)	$F = 11.563$; $p = 0.0044$	[#sk] > [sk#] & [s#k]

cluster to be more overlapped than the heterosyllabic sequence. This can be seen both for overlap and the latency of C2 relative to C1. A consideration of the other heterosyllabic and coda pairs below may illuminate differences between them which did not reach significance here.

The two stop–stop sequences, [g#d] and [gd#] are compared next. The contact profiles for the heterosyllabic sequence [g#d] have already been presented in Section 3.1.—Fig. 3B, D, F, H, and J. The contact profiles for the coda cluster [gd#] are presented for comparison in Fig. 8A–E.

No significant differences in reduction of the consonants in [g#d] *vs.* [gd#] are

TABLE VI. Mean timing measures for [s#k], [#sk], [sk#], pooled across speakers

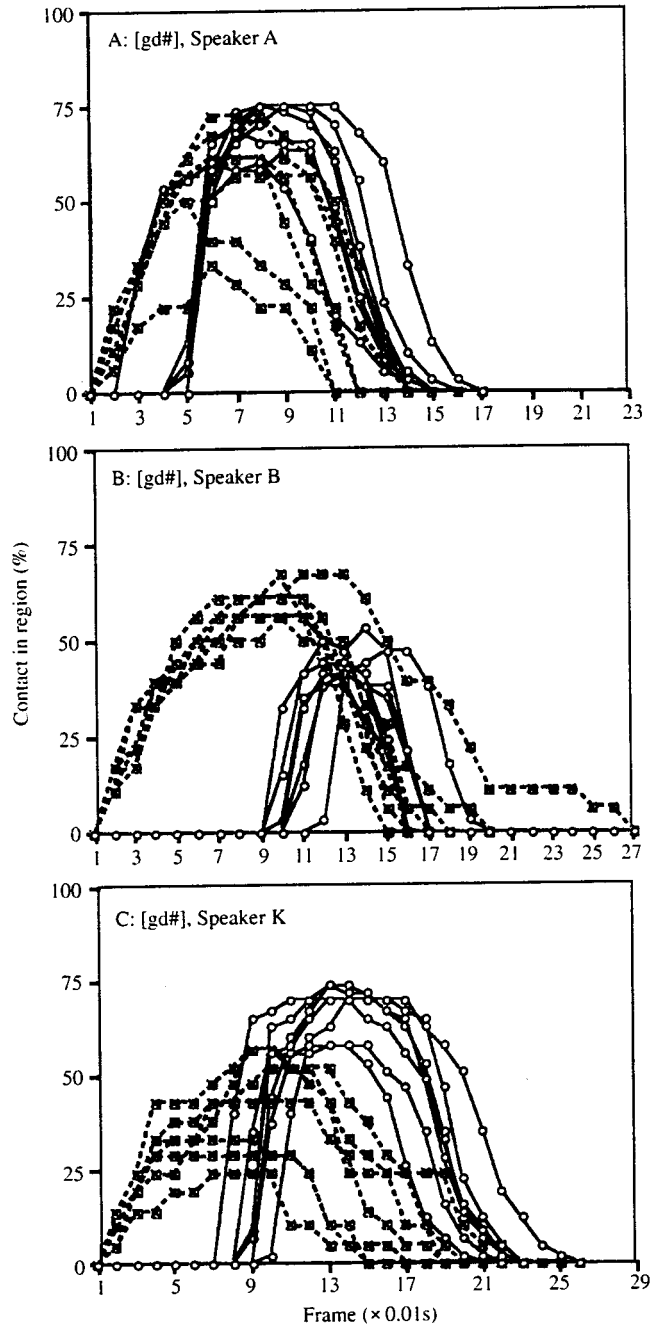| Measure	[#sk]	[s#k]	[sk#]
FRONT DURATION	0.18 s	0.15 s	0.16 s
BACK DURATION	0.12 s	0.13 s	0.14 s
SEQUENCE DURATION	0.22 s	0.20 s	0.21 s
SEQUENCE OVERLAP (%)	36%	41%	44%
C1 OVERLAP (%)	43%	54%	55%
ΔONSETS	0.10 s	0.07 s	0.07 s
ΔPEAKS	0.07 s	0.05 s	0.06 s
ΔC1 RELEASE TO C2 ONSET	0.0 s	−0.02 s	−0.02 s
C2 ONSET RELATIVE TO C1 (%)	56%	46%	44%

Figure 8. Contact profiles for five speakers for [gd#].

observed, although there is a trend ($p \leq 0.08$) for [g] contact to shorten somewhat in the coda cluster. There are some differences in timing though. In accordance with the trend observed for fricative–stop sequences, this coda cluster tends to be more overlapped than the heterosyllabic sequence, as indicated by a trend in C1 OVERLAP
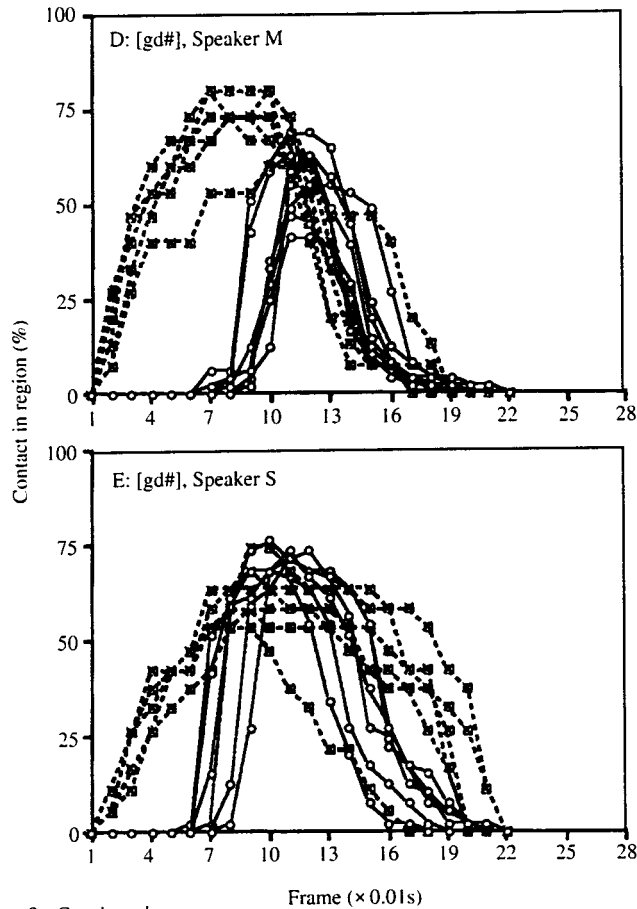
Figure 8. *Continued.*

$(F(1, 3)8.414, p = 0.0625$, Speaker B excluded), and has a shorter ΔONSETS $(F(1, 4) = 11.040, p = 0.0293)$.[6] The coda cluster also has shorter relative latencies than the heterosyllabic sequence as indicated by a trend in C2 ONSET RELATIVE TO C1 $(F(1, 3) = 8.551; p = 0.0613$, Speaker B excluded) and a strong effect on C2 PEAK RELATIVE TO C1 $(F(1, 4) = 246.741, p = 0.0001)$ such that C2 in [gd#] peaks relatively earlier in C1. The pooled mean values for C2 PEAK RELATIVE TO C1 are 67% for [gd#] and 77% for [g#d].

We next consider the stop–fricative sequences to determine if there are differences between [g#s] and [ks#]. The contact profiles for the heterosyllabic sequence [g#s] have been presented in Fig. 5 for Speaker M. The contact profiles for the coda cluster [ks#] are presented for Speaker K and Speaker M in Fig. 9.

---

[6] An opposite effect on ΔC1 RELEASE TO C2 ONSET $(F(1, 4) = 15.211, p = 0.0175)$ ([gd#] > [g#d]) can be explained by noting the difference in the skew of both consonants in the two cases. The C1 ([g]) contact profile in the coda clusters is positively skewed (pooled mean = 0.067), while as a single coda consonant in a heterosyllabic sequence it is negatively skewed (pooled mean = −0.035). This difference in skew brings the C1 release closer to the onset of C2 in the heterosyllabic condition than in the coda cluster. This is presumably responsible for this main effect.
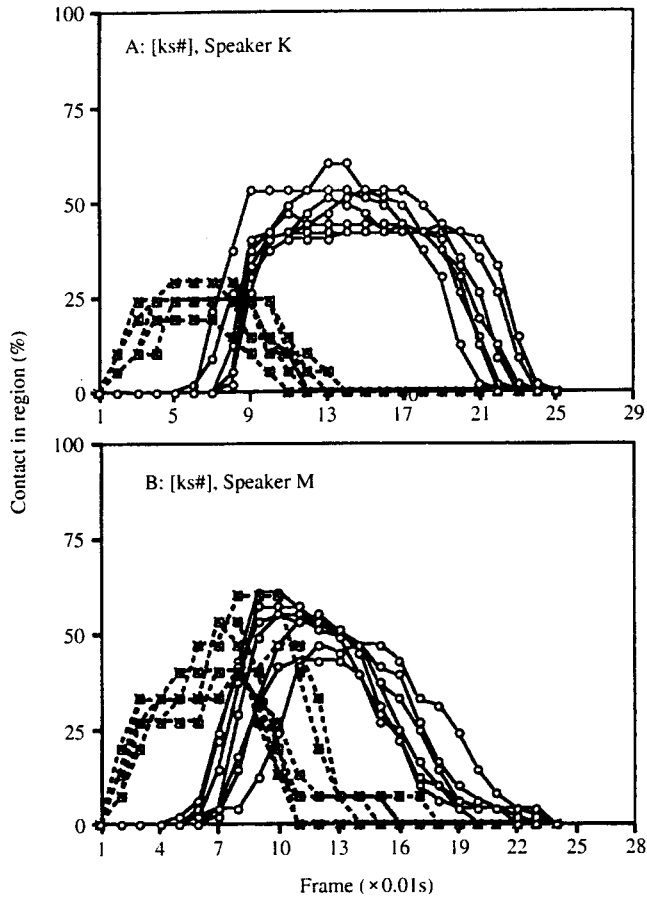
Figure 9. Contact profiles for [ks#] for (A) Speaker K and (B) Speaker M.

We find that the stop–fricative sequences behave differently from the stop–stop sequences. First, there are differences in consonant reduction. The velar stop has less lingua-palatal contact when it is the first member of a coda cluster (34% BACK MAXIMUM) as compared to when it is a single coda consonant (50% BACK MAXIMUM) $(F(1, 4) = 33.765; p = 0.0044)$. This result contrasts with findings for C1 in [s#k] and [sk#] where [s] contact is shorter as single coda consonant than as part of a coda cluster. However, the possibility that the contact differences for the velar in the stop–fricative sequences might be a concomitant of the ancillary voicing difference in this case cannot be ruled out. Finally, the [s] contact is shorter in the coda cluster [ks#] as compared to in the onset position of the heterosyllabic sequence [g#s] $(F(1, 4) = 76.959; p = 0.0009)$. The shortening of the marginal [s] in the coda cluster is different from the behavior of the marginal [d] which does not reduce significantly in the coda cluster [gd#] compared to as a single onset in [g#d]. For the stop–fricative sequences, it appears that both consonants may be subject to reduction in the coda cluster.

The timing effects in the stop–fricative sequences also differ from those observed for the stop–stop sequences. Recall that [gd#] has more overlap and shorter

TABLE VII. Mean overlap, pooled across speakers

	SEQUENCE OVERLAP	C1 OVERLAP
[g#s]—heterosyllabic	27%	50%
[ks#]—coda cluster	20%	38%
[g#d]—heterosyllabic	46%	53%
[gd#]—coda cluster	48%	56%
[s#k]—heterosyllabic	41%	54%
[sk#]—coda cluster	44%	55%
[#sk]—onset cluster	36%	43%

latencies than [g#d]. However, [g#s] and [ks#] show minimal timing differences, and the differences that do exist go in the opposite direction. In [g#s], [s] achieves maximum contact earlier in C1 than it does in [ks#] ($F(1, 4) = 21.556; p = 0.0097$ for C2 PEAK RELATIVE TO C1). Additionally, although not significant differences, all speakers have a greater SEQUENCE OVERLAP for the heterosyllabic sequence, and four of five speakers have greater C1 and C2 OVERLAP in [g#s] and a shorter latency as indicated by C2 ONSET RELATIVE TO C1 for [g#s]. These measures suggest that there is more coproduction in the heterosyllabic stop–fricative cluster than in the parallel coda cluster. However, the entire sequence duration is longer for the heterosyllabic sequence ($F(1, 4) = 27.506; p = 0.0063$).

The difference in duration between the coda cluster [s] and the onset [s] is robust—0.13 s versus 0.16 s. (There is no interaction of [SPEAKER × FRONT DURATION] indicating that the [s] duration behaved similarly for all speakers.) This greater length for the onset consonant is accompanied by greater overlap, although the change in overlap is not enough to compensate for the overall greater sequence duration in the heterosyllabic sequence (0.22 s versus 0.19 s). See Table VII for a summary of *pooled* mean overlap values.

We have seen that the stop–stop, stop–fricative, and fricative–stop sequences behave somewhat differently in terms of their coproduction. Specifically, the coda clusters are more overlapped than the heterosyllabic sequences for the stop–stop sequences, but the reverse tends to be true for the stop–fricative sequences. The fricative–stop sequences have no consistent difference in overlap between the coda clusters and the heterosyllabic sequences.

Finally, consider the timing differences among the *coda* sequences. Theories in which marginal coronals are considered extrasyllabic or as syllable affixes[7] (e.g., Fujimura & Lovins, 1978) might suggest that such a sequence would have less overlap than coda clusters without such an element due to a looser affiliation of the appendix with its syllable. In this case, [ks#] and [gd#] would be likely to be less overlapped and possibly longer than [sk#] due to the marginal coronal consonants.

The coda sequences decrease in total duration from [sk#] to [ks#] to [gd#] ($F(2, 8) = 6.558; p = 0.0206$). The [ks#] sequence is less overlapped than the other two sequences ($F(2, 8) = 9.007; p = 0.0089$). The sequence [gd#] has a shorter latency (as measured by ΔPEAKS) than both of the other sequences ($F(2, 8) = 17.364; p = 0.0012$). The relative latency (as determined by C2 PEAK RELATIVE TO C1)

---

[7] The distributional argument which suggests that [ks#] contains an appendix may well apply to [sk#] as well. Neither occurs as word-medial coda in English; in fact, [ks#] occurs marginally, e.g., [ɛks.tɹə], but [sk#] does not (D. Steriade, p.c.).

decreases, that is, C2 peaks earlier in C1, from [ks#] to [sk#] to [gd#] for all speakers ($F(2, 8) = 29.778$; $p = 0.0002$).

These results suggest that coproduction *decreases* from [gd#] to [sk#] to [ks#]. This is compatible with a prosodic representation postulating marginal [s]'s which are not as tightly bound to the rest of the syllable as the stops considered here. However, there is no evidence for the assumption that marginal [d]'s behave similarly, as would be expected given a theory in which all coronal obstruents are considered to be appendices. The [gd#] sequence has shorter latencies than both [ks#] and [sk#].

Finally, recall that predictions were outlined in Section 1 regarding cohesiveness or variability as a function of syllable structure. Specifically, standard prosodic representations suggest increased overlap and decreased variability for consonants sharing a mora, that is, coda clusters. Additionally, within Articulatory Phonology (Browman & Goldstein, 1990b) it is claimed that "well-formed" sequences will be least overlapped and least variable. These claims will be discussed further in Section 5. The Levene F statistic was calculated in order to evaluate differences in timing variability between onset clusters, coda clusters, and heterosyllabic sequences. Two variables are considered: C2 ONSET RELATIVE TO C1 (relative latency) and ΔC1 RELEASE TO C2 ONSET (absolute latency). All seven sequences—[s#k], [#sk], [sk#], [g#d], [gd#], [g#s], [ks#]—are included and coded as onset cluster, coda cluster, or heterosyllabic sequence.

The results show a significant effect of sequence type on the variability of the relative latency of C2. The onset cluster, in addition to being the least overlapped sequence, is less variable than either coda clusters or heterosyllabic sequences ($F(2, 8) = 4.923$; $p = 0.0404$). The difference in variability between coda clusters and heterosyllabic sequences is negligible.

Next, each set of sequences is tested separately for differences in timing variability between the coda and heterosyllabic sequences. The pairs [g#d] and [gd#], and [g#s] and [ks#], have no significant differences in timing variability. However, the analysis of [s#k], [#sk], and [sk#] determines that speakers, except Speaker B, have more variable relative latencies of C2 in the heterosyllabic sequence [s#k] spanning a word boundary than in the coda cluster [sk#] ($F(2, 6) = 13.762$; $p = 0.0057$, Speaker B excluded). Recall that a consistent difference in degree of overlap is not observed in these two sequences.

4. Discussion

4.1. *Place, manner, and order effects*

4.1.1. *Coda reduction*

In summary, analyses of the heterosyllabic sequences [g#d], [d#g], [g#s], and [s#g] demonstrate the following with respect to coda reduction:

- In the stop–stop sequences, the stop articulation has less lingua-palatal contact in coda position than in onset position.
- In the stop–fricative versus fricative–stop sequences, lingua-palatal contact for the velar stop is smaller and tends to be shorter in coda position than in onset position.

• In the stop–stop sequences, stops in coda position do not significantly differ in contact duration from those in onset position.

These results generally support H1 regarding coda reduction—that there is less contact for a consonant in coda position than in onset position. However, its support extends mainly to the stop consonants. Sequences involving the alveolar fricative show only a small tendency for the fricative to have less and shorter contact in coda position as compared to onset position. Maximum contact in the back region for [g] was less for [g#C] (56%) than for [C#g] (71%) ($F(1, 3) = 20.489$ $p = 0.0202$). The effect on the front consonant was similar but stronger for the stop than for the fricative. While the method employed here does not allow a direct comparison between tongue tip and body in the degree of reduction, Browman & Goldstein (1992b) and Browman (1994) have suggested that a greater magnitude of reduction in word-final coronals may be a result of the tendency of the tongue tip to rest on the floor of the mouth during vowels thereby requiring it to travel further than the tongue body or lips to form a closure. The contact differences observed here between codas and onsets accord with other findings reported on coda reduction. Word-initial [b]'s have been found to have both longer and greater muscle activity associated with them as compared to word-final [b]'s (Fromkin, 1965). Other experiments have reported articulatory data showing reduction of both singleton consonants in coda position (e.g., Browman & Goldstein, 1995) and coda consonants in clusters (e.g., Barry, 1985, 1991; Kerswill, 1985; Nolan, 1992).

With respect to variability and reduction, analyses of the heterosyllabic sequences [g#d], [d#g], [g#s], and [s#g] demonstrate the following:

• At least for alveolar consonants, the same dimensions that are subject to reduction in coda position are generally more variable in coda position than in onset position.

This is relevant to H2—that lingua-palatal contact is more variable in coda position than in onset position. However, it does not specifically support this more generally stated hypothesis. In [d#g] and [g#d], maximum [d] contact is reduced and more variable in coda position. There is no effect of syllable position on the means or the variability of the duration of [d] contact, or the maximum contact for [g]. However, there is a significant effect on the variability of duration of lingua-palatal contact for [g] (greater variability for coda [g] in [g#d] than onset in [d#g]), although there is no significant main effect on the mean durations. In the analysis of [g#s] and [s#g], there are also parallels between reduction and variability in the front consonant, in this case a lack of both. The fricative, which does not reduce in maximum contact when a coda, also does not show increased variabiilty in contact in this position. With respect to the stop/fricative asymmetry in variability, MacNeilage (1970) outlines a suggestion originating with Stevens & House (1963): "It may . . . be that targets for fricatives are specified with more precision than targets for stop consonants because the acoustic result is more dependent on precise articulator position in the former case" (p. 193). Reduction in magnitude is, however, indicated for [g] in coda position in this pair of sequences, but it is accompanied by no consistent difference in variability. Lastly, the exceptional behavior of Speaker B might be taken as support for the compatibility of reduction and variability. Speaker B, who does not exhibit much coda reduction of [d] in [d#g], also does not have

more variability in maximum displacement in coda position. In fact, her onset [d]'s have relatively less contact and are more variable. Perhaps Speaker B's exceptional data are due to a more careful style of speech.[8] In summary, the co-occurrence of reduction and variability in coda coronals suggests that variability may play an important role in reduction, at least for the alveolar consonants.[9]

4.1.2. *Place and manner differences in sequence timing*

With respect to the timing of the consonant sequences, the following place and manner effects were observed in the analysis of the heterosyllabic sequences [g#d], [d#g], [g#s], and [s#g]:

- Lingua-palatal contact profiles show tongue tip consonants to be more overlapped by a following tongue body consonant than a tongue body consonant is by following tongue tip consonants.
- Contact profiles show an alveolar fricative to be less overlapped than an alveolar stop by a following velar stop. More generally, sequences involving a fricative are less overlapped than those which only have stops.

These two effects support H3 and H4, respectively. SEQUENCE OVERLAP is significantly greater in the front-back pair ($F(1, 4) = 10.922$, $p = 0.0298$), and the tongue tip consonants are significantly more overlapped by a following [g] than the reverse ($F(1, 4) = 22.244$, $p = 0.0092$).

Catford (1977:222), using EPG, estimates the degree of articulatory overlap between two adjacent consonants as *ranging* from 29% to 45% of the total sequence duration. However, he does not report the number of subjects studied or the number of repetitions. We observed here *means* of between 27% and 59% overlap of contact for two consonants in sequence. If Catford was using total contact duration to indicate sequence duration, which does not seem unreasonable, we conclude that we observe generally more overlap than he. We also observe a much wider *range* of overlap, with sequences ranging from 11% to 91% overlapped.

We also saw that [d] in either C1 or C2 position was more overlapped than [s] in the same position. Why should this be? There may be a perceptual motivation for this difference.[10] The abrupt discontinuities at the edges of a stop consonant are important perceptual cues in recovering that consonant (Stevens & Keyser, 1989). For a fricative, however, abrupt discontinuities or edges are not as important as the frequency distribution of the fricative noise itself. (Stevens & Keyser (1989) note the importance of the acoustic consequences of a slower fricative release.) It may be that there is a perceptual motivation for not obscuring too much of the [s] duration

---

[8] This speaker is both older than the others and, due to a background in singing, had undergone pronunciation training in the past. These may be factors encouraging a tendency to produce more careful, formal, or slower speech. (On such an age-related effect, see Welford, 1977; Kent & Burkard, 1981; Ramig, 1983; Chodzko-Zajko & Offenback, 1986; and Amerman & Parnell, 1992).

[9] The relationship between spatial displacement and variability is one of the issues inspiring Keating's "Window model" of coarticulation (Keating, 1990). To account for our displacement data in the "Window model", onset stops would have to have more specific targets, that is, more narrow windows, than coda stops. Importantly, to account for our data, the window would have to have a lower *bottom* margin, not simply be displaced downward. Note, however, that Keating's Window model loses some of its appeal if the target projection from a segment's features must take place in a context-dependent fashion. As the prosodic affiliation of a segment is not specified featurally, window projection by (certain) features would have to have access to prosodic information, a substantial empowerment.

[10] I thank Ian Maddieson for suggesting this issue to me.

with a coarticulated stop. However, for [d] only an edge need be discernible as the abrupt change in amplitude is the perceptually salient cue. (See also Jongman, 1989.) This would create a situation in which it is important not to overlap too much of [s], while [d] can withstand being more overlapped. The perceptual product and articulatory constraints are not independent. To put the same idea another way, the back of the tongue mustn't raise too early when [s] precedes [g], as this posture would prevent the appropriate airflow through the alveolar constriction thereby impeding the production of the fricative. Speakers avoid doing this thereby ensuring that the fricative will have the acoustic cues which in turn allow it to be accurately recovered by the listener. (In fact, for an [s] to be identifiable it must have approximately 50 ms of frication (Jongman, 1989). For these [s#g] sequences, the mean time between the *first* maximum contact for [s] and that for [g], a rather rough approximation of the frication noise duration, is 53 ms, suggesting that they might be overlapped as much as possible without impeding identifiability.) Hardcastle (1985) reaches similar conclusions in examining coarticulation in /kl/ sequences, stating that variation in overlap of /l/ during the /k/ closure will have little perceptual effect. He sees the principle constraint on the degree of coarticulation as perceptual.

Ladefoged, DeClerk, Lindau & Papçun (1972) and Johnson, Ladefoged & Lindau (1993) have outlined an auditory theory of speech production in which „Lindau (1993) have outlined an auditory theory of speech production in which speech movements are directed by auditory goals and have gone on to suggest that "the acoustic product of speaking is the crucial determinant of the *organization* of speech articulation" (Johnson *et al.,* 1993, emphasis added, p. 713). To the extent that a model of speech organization must allow for acoustic and perceptual goals to be *an* influence on articulatory timing, this concept is profitable. The difference in [s] and [d] overlap serves as an example of how this type of influence may operate. Browman & Goldstein, while not incorporating acoustic influences into Articulatory Phonology, do acknowledge that "differential acoustic effects may lead to a preference for one kind of [intergestural] organization over the other . . ." for use in linguistic contrast (Browman & Goldstein, 1992b, p. 226). I suggest that such influences may also play a role in determining inter-gestural phasing relations not specific to a linguistic contrast, such as the across-word-boundary timing examined here. (For an example of how a variety of different influences on phasing relations could be incorporated into an Articulatory Phonology-like framework, see Byrd, in press.)

The question of why the front-back sequences are more overlapped than the back-front sequences also bears on this point. Paralleling the findings here, Hardcastle & Roach (1979) found that the time between the initiation of closure for an adjacent [t] and [k] was shorter in a [tk] cluster than for a [kt] cluster, at least in a front vowel environment. They concluded that this was because the movement from [t] to [k] involves the contraction of a single intrinsic tongue muscle to raise the back of the tongue while the movement from [k] to [t] requires the use of two muscles including the extrinsic genioglossus to reposition the tongue upwards and forwards. Recasens *et al.* (1993) suggest that one reason for the difference found by Hardcastle & Roach could be that the tongue tip has a greater "degree of flexibility" and that this causes greater anticipatory coarticulation, although the exact connection between these assertions is not clear. In our data, the rather large difference observed between [d#g] and [g#d], with contact for the consonants often starting nearly simultaneously in [d#g], does not seem to support Hardcastle & Roach's

explanation. The difference in time taken to contract one intrinsic muscle versus one extrinsic and one intrinsic muscle is likely to be incredibly short. In fact, it is not even clear that the assumption that using two muscles takes longer than using one is valid. The well-supported conception of muscle groups organized into coordinative structures suggests that it is unlikely that the exact number of muscles involved should create differences in the time between articulations. In order to describe the large differences in temporal latency for [d#g] and [g#d], some difference in interarticulator timing, not simply execution, may be required.

So, why should [d#g] be more overlapped than [g#d]? Speakers may make less of an effort to preserve less robust perceptual cues.[11] Because the cues for the unreleased [d] are so weak to start with, there is little motivation for the speaker to safeguard them. Even a little adulteration from C2 can overwhelm them. The formant transitions for [d] show relatively small excursions (Öhman, 1967; Lade-foged, 1993) and don't have a large effect on F_3. Also, [d] has an articulation, and hence, formant movement, that is relatively rapid (Kuehn & Moll, 1976). Lastly, final alveolars in VC contexts are perceptually more confusable than bilabials and velars (Winitz, Scheib, & Reeds, 1972). The slight benefit which would be gained by lessening the overlap for [d#g] might not be worth the cost of the loss of transmission speed. (One of the most important aspects of communicative efficiency is parallel transmission, whereby information about several linguistic units is transmitted simultaneously in tandem. See Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967 and Mattingly, 1981.) As the listener will presumably be able to tell that two consonants were present due to closure duration and top-down knowledge, [d] may well be the default "guess" when there are no strong acoustic cues present suggesting an alternative. The other obvious possibility is that the listener may assume a geminated C2, thereby giving rise to the phenomena of place assimilation common in [dC] clusters.

4.2. Sequence type and timing

We have found evidence that onset clusters, coda clusters, and heterosyllabic sequences differ in their inter-gestural timing and in reduction. However, the precise nature of these effects depends on the consonants in the sequence. The onset cluster investigated here is less overlapped and less variable in its timing than the coda clusters and heterosyllabic sequences. This parallels results reported by Hardcastle (1985) for /#kl/ *vs.* /k#l/ sequences who found less overlap for the onset cluster. For the fricative–stop sequences, minimal differences in the timing of coda clusters and heterosyllabic sequences are found. In the stop–stop sequences, coproduction is greater in the coda clusters than in the heterosyllabic sequences. In the stop–fricative sequences, however, there is some evidence that the coda cluster is less

[11] A similar line of reasoning has been advanced by Kohler (1992) with respect to reduction. Regarding coda reduction, he says "[w]hat is not very distinctive for a listener anyway may be reduced by a speaker more easily to yield to the principle of economy of effort" (p. 231). Browman & Goldstein (1992b) respond that "initial gestures would have higher signalling value to listeners than final ones because they are reduced less (not vice versa)" (p. 231). A similar response could be made to our suggestion that relatively less perceptually valuable cues are more likely to be obscured by temporal *overlap*. However, low-level articulator based effects probably structure words to have more distinctive information in certain positions, to which listeners have learned to attend; in turn, speakers take this into consideration in deciding how to allocate articulatory resources in production (P. Keating, p.c.).

overlapped than the heterosyllabic sequence. In general, these findings do not support H5, that overlap is greater when clusters are tautosyllabic. Also, H5′, that the coda clusters, forming a moraic constituent, exhibit more overlap, receives only minimal support.

As an aside, consider again that the [#sk] onset cluster is less overlapped than the coda and heterosyllabic clusters. If this result were replicated for other /sC/ clusters, a hint regarding an interesting morpheme structure constraint in English is suggested. In English, homorganic non-coronals may not flank both sides of the vowel in /sCVC/ sequences (Davis, 1990). Tongue tip consonants of any type may appear in this position, however. Davis (1990) points out that this is not a constraint against identical consonants but rather against consonants having the same oral place of articulation; nasality, for example, is irrelevant.[12] If /sC/ onset clusters regularly have little overlap, this could yield a situation in which the C's in a /sCVC/ sequence are "pushed" close together, requiring a rapid opening and re-closing of the articulator. Such a structure may be disfavored. However, in the case of /sCVC/ sequences where only a single articulatory subsystem is required for /sC/, i.e., the tongue tip, the situation is one of gestural *blending* (Browman & Goldstein, 1990a and elsewhere) rather than overlap between gestures using different articulators. This might allow greater coproduction in such sequences than in /#sk/ or /#sp/. Additionally, if the tongue tip is significantly faster than the other articulators, this could also facilitate the articulation of the second C in /sCVC/ sequences in which C is coronal. These factors might motivate the exceptionality of coronals to this morpheme structure constraint. The constraint itself might exist in response to the relatively small amount of overlap in /sC/ onsets. Additionally, with this approach, nasality is predicted, correctly, to be irrelevant to the constraint since only the oral articulations are at issue.

4.2.1. *Association and well-formedness*

Next, consider Browman & Goldstein's (1990b) account of timing in consonant sequences, which bears both on degree of overlap and variability in overlap. Within this framework "associated" consonants are phased with respect to one another by a rule synchronizing the onset of C2 to the offset of C1. The following is a proposed statement of association for consonant sequences:

> The leftmost consonantal gesture of a consonant sequence intervening between two vocalic gestures is associated with both vocalic gestures. A consonant sequence is defined as intervening iff the entire sequence lies between the two vocalic gestures when projected onto the one-dimensional oral tier. (Browman & Goldstein, 1990b: 257)

Browman & Goldstein use a heterosyllabic sequence in an example of this statement's application, associating the coda C to the vowel of the following word. After reassociation, the rule for consonant cluster phasing then operates if the sequence is "well-formed." A sequence is well-formed if its gestures on the C-tier (i.e., oral gestures) form possible onsets (or codas) if reassociated to the following vowel. (As it is the *first* consonant which is associated to both vowels, the status of

---

[12] Davis (1990) reports that *skunk* is the only exceptional word found in a 20,000-word database and notes that *spam, skag,* and *spumoni* would also be exceptional had they been in the database.

possible onset seems to be the crucial one, Browman & Goldstein, 1990b, also citing Kent & Moll, 1975, to this effect). However, no explicit differentiation between possible onset and possible coda is made in the rule formulation or discussion.) If the sequence is not well-formed, the rule does not operate and the consonants are not phased with respect to one another. In such cases, Browman & Goldstein (1990b) make two predictions. First, such sequences should allow massive, some-times complete, gestural overlap which may obscure some consonants. Second, such sequences do not have the "same kind of tight organization available", and for such sequences "variation in the degree of overlap is possible" (Browman & Goldstein, 1990b, p. 369).

In this experiment, the timing measures generally are not significantly different in variability among the heterosyllabic sequences, the one exception being the significantly greater variability of [g#s] as compared to [s#g] in the relative latency of C2. This is relevant to the claim of Browman & Goldstein (1990b) regarding the relationship between overlap and the status of the sequence as a possible onset or coda. Consider Browman & Goldstein's predictions regarding overlap and variabi-lity as a function of "resyllabifiability". If a well-formed sequence consists of oral gestures that form either an allowable onset *or* an allowable coda, then of the set [g#d], [d#g], [g#s], and [s#g], only [d#g] is excluded. In fact, given this criterion of well-formedness, it is difficult to find any two-member heterorganic sequence in English, other than those having an alveolar stop (/t/, /d/, /n/) in first position, whose oral gestures do not comprise a possible onset or coda. (See Prator & Robinett (1985) who list 38 two-member onset sequences and 65 two-member coda sequences.) Indeed, [d#g] is the most overlapped sequence for all speakers and the most variable in relative timing for four of the five speakers. This is in accordance with Browman & Goldstein's claims. However, the variability is not significantly different from [g#d]. It could be that the criterion of possible onset or coda acts to predict increased overlap and variability exactly, and only, when [d] or [t] is non-final in a cluster. It is possible that in English the increased coproduction of such sequences is due to the nature of C1 rather than the well-formedness of the sequence as a whole. A relevant test case might be a [tr] sequence where the [t] is initial but in an acceptable cluster.

Alternatively, if Browman & Goldstein's criteria for well-formedness refer only to a possible *onset,* only [s#g] of these four sequences has oral gestures which form a possible onset. Indeed, [s#g] is significantly less variable (for four speakers) than [g#s]. It is, however, significantly *more* overlapped. Thus, Browman & Goldstein's prediction succeeds on one count but fails on the other.

For the three sequence types examined here, Browman & Goldstein's (1990b) reassociation process described above would operate to yield the associations shown in Fig. 10. The phasing associations between the C and V tiers added by the application of this statement are shown by heavy lines; underlying associations are shown by lighter lines. (Vacuous application of the association statement is not shown.) Note that, within a tier, canonical sequencing relations are preserved visually, and between tiers, the sequencing is conveyed by the angle of the phasing association lines.

However, Browman & Goldstein (1990b) suggest that a final [s] gesture does not co-occur, that is, is not associated, with the preceding vowel. Following this suggestion, the reassociation of the [ks#] cluster yields the result shown in Fig. 11.
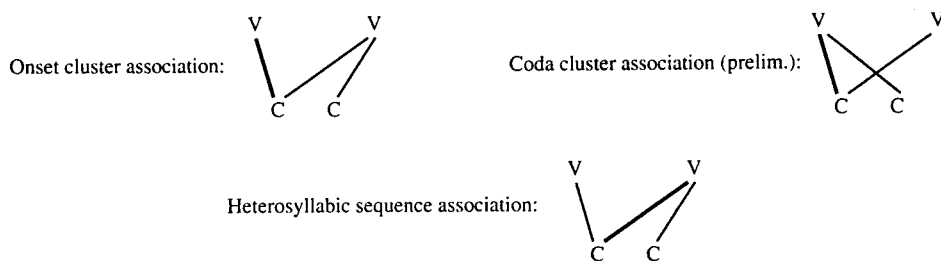
Onset cluster association: Coda cluster association (prelim.):

Heterosyllabic sequence association:

Figure 10. Phasing associations of C and V tiers.

Browman & Goldstein (1990b) do not elaborate on the association of consonants other than [s] in this position. Their consonant cluster phasing rule synchronizing the onset of C2 with the offset of C1 should apply only to consonants that are associated to the same vowel. Coda clusters with a marginal [s] should not undergo the rule.

As can be seen, identical association relationships result for onset and heterosyllabic clusters. Therefore, only the application, or lack thereof, of the phasing rule should engender timing differences. Whether the rule applies or not depends on the well-formedness of the sequence. If the status of possible onset *or* coda is the significant determiner of its application, then the onset and heterosyllabic sequences examined here are predicted to be timed similarly (and consistently). This prediction is not supported. If being a possible onset is the *single* determiner of the application of this phasing rule, then the timing of the [#sk] onset should be different from the other sequences. To a certain extent this was the case. It was less overlapped and more stable, as predicted. However, it should also be the case, all else being equal, that [g#s] and [ks#] behave similarly. Neither should undergo the phasing rule—the former because it is not a potential onset and the latter because it is not properly associated. This was not the case; timing differences were observed here. Also, these sequences were the least overlapped. This is not in accordance with Browman & Goldstein's prediction of large overlap when the CC Phasing Rule fails to apply. It is not clear if differences are predicted to exist between [g#d] and [gd#] and between [s#k] and [sk#] since we aren't certain about the predicted association of the coda cluster consonants. The association and phasing statements do not illuminate the timing differences *between* the various coda clusters or the differences in timing and variability found between the coda clusters and the heterosyllabic sequences. They do account nicely for the behavior of the onset sequence in comparison to the other sequence types.

4.2.2. *Syllable structure representation and timing*

There are three questions of interest here regarding the standard phonological representations of clusters. First, consider Selkirk's (1982) proposal that s + obstruent onset and coda clusters be considered single segments (see also Steriade, 1982 and Lamontagne, 1993). Browman & Goldstein (1990b) suggest that the
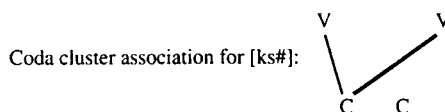

Coda cluster association for [ks#]:

Figure 11. Coda cluster phasing association (*revised*).

presence of two oral constrictions and one glottal constriction for s-stop sequences allows the possibility of their acting as single units or as two units. Our observation of relatively less overlap in the onset [#sk] seems to run counter to other cases where other pairs of (non-laryngeal) constrictions form a single segment. Additionally, within the examination of the coda clusters, the [sk#] coda cluster was less overlapped than the [gd#] cluster. These findings suggest that the supposition of s-obstruent clusters being a single segment in contrast to other clusters does not relate very transparently to intuitive notions regarding the greater extent of overlap within segments than across them.[13] However, the stability of the [#sk] timing should be noted, as it has been hypothesized (Byrd, 1995b) that timing stability may play a defining role in the traditional notion of segmenthood.

Secondly, there was only marginal evidence (4 speakers, 1 sequence type) for decreased variability in coda clusters relative to heterosyllabic sequences due to greater gestural cohesion, as is suggested by the H6'. In most cases, timing variability observed in coda clusters is like that in heterosyllabic sequences which do not comprise a word-level superordinate constituent. Onset clusters on the other hand exhibit less variability and less overlap than coda clusters or heterosyllabic sequences. This can only be a tentative conclusion as [#sk] is the single onset sequence here. The effect was robust, however. Given the assumptions outlined in Section 1 regarding the relationship between constituent structure and timing and stability, the relative stability of the onset cluster is not predicted by the standard constituent structure representation in which consonants in onset clusters form no single constituent, unlike those in coda clusters.

Because this experiment was restricted to two member sequences, resyllabification has not been directly controlled for experimentally. This restriction on the design yields stimuli in which the coda clusters are prevocalic. However, the clusters do *not* precede an unstressed syllable but rather a syllable receiving about the same level of stress as the syllable in which the cluster occurs. This should lessen the likelihood of resyllabification.[14] Even so, possible resyllabification of the cluster cannot be ruled out, especially in light of the small degree of difference observed between the coda clusters and the heterosyllabic sequences.

Lastly, recall the standard representation of the coda cluster in which the coda consonants share a single mora, as shown in Fig. 1. There is no princple of moraic

---

[13] One of the reasons that s-stop clusters are sometimes considered to form a single segment is the disinclination of such sequences to allow epenthesis. One reason for this may be that an unusually abducted glottis is found for /s/'s in comparison to voiceless stops due to the necessity for a high rate of airflow (Hirose & Gay, 1972; Hirose, Lisker & Abramson, 1972; Collier, Lisker, Hirose, & Ushijima, 1979; and Yoshioka, Löfqvist, & Hirose, 1982). Additionally, in a number of languages only one glottal gesture is found for words beginning with s-stop clusters (Petursson, 1977; Löfqvist & Yoshioka 1980a, 1981b; Browman & Goldstein, 1986 and Goldstein, 1990 citing Yoshioka, Löfqvist, & Hirose, 1981; and Fukui & Hirose, 1983). Because the stop in s + stop sequences is generally voiceless, the insertion of an epenthetic vowel would require a rapid change from a very abducted glottis to a brief moment of voicing for the epenthetic vowel to, again, an abducted glottis for the voiceless stop. (I thank Richard Wright for pointing out this possibility to me.) This would disfavor such epenthesis. Other clusters of voiceless obstruents may have multiple openings of the glottis (Kingston, 1990 citing Löfqvist & Yoshioka 1981a,b and Yoshioka, Löfqvist, & Hirose, 1981, 1982). These openings could also be less extreme when [s] is not involved.

[14] Note especially that the coda sequence with the *least* overlap, [ks#], was also the sequence that was followed by the largest phonological phrase boundary: "Say backs Abigail."
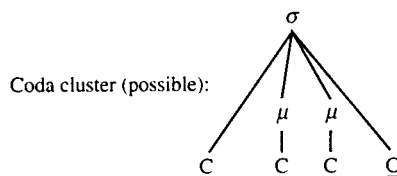
Figure 12. A possible coda cluster representation.

theory which requires this to be the case. A possible representation in which only the first of the coda consonants is moraic, as shown in Fig. 12, can also be entertained.

This representation is possible in principle, is more compatible with the findings above, and is suggested by independent evidence. Such a representation creates a greater parallelism between the coda cluster structure and that of the heterosyllabic sequence. This is in line with our data showing no consistent timing differences between these two types of sequences. It also seems in keeping with the lack of magnitude reduction observed for C2 in the coda cluster as compared to that consonant in onset position, suggesting parallelism between the non-moraic status of C2 in a coda cluster and in an onset. Furthermore, consonant duration facts in Swedish (Löfstedt, 1992) lend independent support for the possibility of such a structure.[15] Löfstedt describes the first consonant of a (monomorphemic) coda cluster, and only this consonant, as long under stress. (See also Prince, 1980 on a similar phenomenon in Estonian.) This suggests that C1 is different from C2 in a coda cluster with respect to its rhythmic, that is, moraic, properties in the language. This might be the case in English as well. These facts regarding overlap, reduction, and duration are all compatible with the alternative representation of coda clusters shown in Fig. 12 in which the first consonant of the complex coda is moraic and the following consonant is directly dominated by the syllable.

In summary, no consistent evidence of overlap or variability differences between tautosyllabic and heterosyllabic sequences is found. Thus, H5 and H6 are not supported. Additionally, the data examined do not appear compatible with the prosodic representations shown in Fig. 1 given the assumptions outlined as to the relation between constituency and articulatory timing. The representation shown in Fig. 12 is suggested as a possible improvement.

5. Conclusion

This work investigated the nature of certain linguistic influences on the articulatory timing of two consonants in sequence. Degree of temporal overlap and variability in timing were both considered. Consonantal place, manner, and sequence syllabification were all demonstrated to have an effect on consonant sequence timing. These findings, as well as those regarding timing variability, have implications for an understanding of consonant sequence syllabification. Additionally, consonantal reduction was evaluated in terms of lingua-palatal contact degree and duration and was found to vary as a function of place, manner, and syllable position. Other work (Byrd & Tan, this issue) has shown [C#C] timing to be affected by speech rate.

---

[15] I thank Donca Steriade for pointing the Swedish data out to me.

The results suggest that Articulatory Phonology is an interesting framework in which to consider inter-gestural timing. Consonant sequence timing was found to be variable and influenced by a number of factors. Specifically, /CC/ overlap was found to be, to some degree, lawful in its variation. Timing variability is not illuminated by Articulatory Phonology's current rule-based conception of phasing relations, and these results encourage speculation as to how it might be conceptualized.

Keating (1995) sees variation in phasing as a function of prosody, especially within a word, as a potentially fatal shortcoming of Articulatory Phonology. However, at least one way has been suggested within Articulatory Phonology of addressing the problem of invariant, lexically-specified phasing relations. (In fact, this approach is inspired in part by Keating's own Window model of coarticulation (Keating, 1990), a segmental targets-and-interpolation model that posits target ranges.) In work related to this experiment, Byrd (in press) outlines a proposal called the PHASE WINDOW framework in which a PHASE WINDOW constrains the permissible inter-gestural timing relations. Competing linguistic and extra-linguistic influences that differ from utterance to utterance weight a PHASE WINDOW, determining where in the range of permissible overlap relationships a token will actually be realized. For example, consonant clusters would have a PHASE WINDOW which includes $C2(0°) = C1(290°)$, the relationship proposed by Browman & Goldstein (1990b), but are not limited to only that relationship. Such phasing variability, or wide PHASE WINDOWS, are suggested to be implemented post-lexically for /CC/, /VC/, and /CV/ timing relationships. Additionally, Byrd (in press) hypothesizes that a limited number of inter-gestural phasing relations are lexically specified and stable, *i.e.*, have narrow PHASE WINDOWS that are part of their lexical representation. Furthermore, it is suggested that these gestures whose coordination, or PHASE WINDOW, is part of the word's lexical representation, bear a close relation to those conglomerates of gestures that constitute what is traditionally considered to be a "segment" (Byrd, in press).

In conclusion, the experimental results suggest that postulation of invariant phasing rules as the mechanism for implementing inter-gestural timing in Articulatory Phonology must be revised to allow for linguistic and extra-linguistic influences on certain phasing relations. Most importantly, further experiments investigating factors affecting articulatory timing, particularly those that are prosodic in nature, are necessary in order to understand not only the speech production mechanism but its relationship to complex linguistic structure.

References

Abacus Concepts (1989) *SuperAnova*. Abacus Concepts, Inc., Berkeley, CA.
Abercrombie, D. (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
Amerman, J. D. & Parnell, M. M. (1992) Speech timing in the elderly, *Journal of Phonetics*, **20**, 65–76.

Barry, M. (1985) A palatographic study of connected speech process. *Cambridge Papers in Phonetics and Experimental Linguistics,* **4,** 1–16.

Barry, M. (1991) Temporal modelling of gestures in articulatory assimilation. In *Proceedings of the XIIth international congress of phonetic sciences,* **4,** 14–17.

Barry, M. (1992) Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication,* **11,** 393–400.

Beckman, M. E., Edwards, J. & Fletcher, J. (1992) Prosodic structure and tempo in a sonority model of articulatory dynamics. In *Papers in laboratory phonology II: gesture, segment, prosody* (G. J. Docherty & D. R. Ladd, editors). Cambridge: Cambridge University Press, 68–86.

Browman, C. P. (1984) Lip aperture and consonant releases. In *Papers in laboratory phonology III: phonological structure and phonetic form* (P. Keating, editor), pp. 331–353. Cambridge: Cambridge University Press.

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

Browman, C. P. & Goldstein, L. (1988) Some notes on syllable structure in articulatory phonology. *Phonetica,* **45,** 140–155.

Browman, C. P. & Goldstein, L. (1989) Articulatory gestures as phonological units. *Phonology,* **6,** 201–252.

Browman, C. P. & Goldstein, L. (1990a) Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics,* **18,** 299–320.

Browman, C. P. & Goldstein, L. (1990b) Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology I: between the grammar and the physics of speech* (J. Kingston & M. E. Beckman, editors), pp. 341–376. Cambridge: Cambridge University Press.

Browman, C. P. & Goldstein, L. (1991) Gestural structures: distinctiveness, phonological processes, and historical change. In *Modularity and the motor theory of speech perception* (I. Mattingly & M. Studdert-Kennedy, editors), pp. 313–338. New Jersey: Lawrence Erlbaum Associates.

Browman, C. P. & Goldstein, L. (1992a) Articulatory phonology: an overview. *Phonetica,* **49,** 155–180.

Browman, C. P. & Goldstein, L. (1992b) Response to commentaries. *Phonetica,* **49,** 222–234.

Browman, C. P. & Goldstein, L. (1995) Gestural syllable position effects in American English. In *Producing speech: contemporary issues for Katherine Safford Harris.* (F. Bell-Berti & L. Raphael, editors). Woodbury, NY: AIP Press.

Byrd, D. (1992) Perception of assimilation in consonant clusters: a gestural model. *Phonetica,* **49,** 1–24.

Byrd, D. (1994) *Articulatory timing in English consonant sequences.* UCLA dissertation. In *UCLA Working Papers in Phonetics,* **86.**

Byrd, D. (1995a) Articulatory realization of single and blended lingual gestures. In *Proceedings of the XIIIth international congress of phonetic sciences* (K. Elenius & P. Branderud, editors), pp. 438–441.

Byrd, D. (in press) A Phase Window framework for articulatory timing. *Phonology.*

Byrd, D. (1995b) C-Centers revisited. *Phonetica,* **52,** 285–306.

Byrd, D., Flemming, E., Mueller, C. A. & Tan, C. C. (1995) Using regions and indices in EPG data reduction. *Journal of Speech and Hearing Research,* **38,** 821–827.

Byrd, D. & Tan, C. C. (this issue) Saying consonant clusters quickly. *Journal of Phonetics,* **24.**

Catford, J. C. (1977) *Fundamental problems in phonetics.* Bloomington: Indiana University Press.

Choi, J. D. (1992) *Phonetic underspecification and target interpolation: An acoustic study of Marshallese vowel allophony.* Ph.D. dissertation, UCLA. *UCLA Working Papers in Phonetics,* **82.**

Collier, R., Lisker, L., Hirose, H. & Ushijima, T. (1979) Voicing in intervocalic stops and fricatives in Dutch. *Journal of Phonetics,* **7,** 357–373.

Davis, (1990) An argument for the underspecification of [Coronal] in English. *Linguistic Inquiry* **21**(2), 301–305.

Dixon, W. J. (1988) *BMPD statistical software manual.* The Regents of the University of California.

Fromkin, V. (1965) *Some phonetic specifications of linguistic units: an electromyographic investigation.* Ph.D. diss., UCLA. *UCLA Working Papers in Phonetics,* **3.**

Fujimura, O. & Lovins, J. B. (1978) Syllables as concatenative phonetic units. In *Syllables and segments* (A. Bell & J. B. Hooper, editors), pp. 107–120. New York: Elsevier North-Holland.

Fukui, N. & Hirose, H. (1983) Laryngeal adjustments in Danish voiceless obstruent production. *Annual Bulletin Research Institute of Logopedics and Phoniatrics,* **17,** 61–71.

Goldstein, L. (1990) On articulatory binding: comments on Kingston's paper. In *Papers in laboratory phonology I: between the grammar and physics of speech* (J. Kingston & M. Beckman, editors). pp. 445–450.

Hardcastle, W. J. (1985) Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication,* **4,** 247–263.

Hardcastle, W. J. & Roach, P. (1979) An instrumental investigation of coarticulation in stop consonant sequences. In *Current issues in the phonetic sciences* (H. Hollien & P. Hollien, eds.), pp. 531–540. Amsterdam: John Benjamins.

Hayes, B. (1989) Compensatory lengthening in moraic phonology, *Linguistic Inquiry*, **20**, 253–306.

Henderson, J. B. & Repp, B. (1982) Is a stop consonant released when followed by another stop consonant? *Phonetica*, **39**, 71–82.

Hirose, H. & Gay, T. (1972) The activity of the intrinsic laryngeal muscles in voicing control, *Phonetica*, **25**, 140–164.

Hirose, H., Lisker, L. & Abramson, A. (1972) Physiological aspects of certain laryngeal features in speech production, *Haskins Laboratories Status Report on Speech Research*, **Sr-31/32**, 183–191.

Johnson, K., Ladefoged, P. & Lindau, M. (1993) Individual differences in vowel production, *Journal of the Acoustical Society of America*, 94(2), pt. 1, 701–714.

Jones, D. (1956) *An outline of English phonetics, eighth edition* (Eighth Edition). Cambridge: W. Heffer & Sons.

Jongman, A. (1989) Duration of frication noise required for identification of English fricatives, *Journal of the Acoustical Society of America*, 85(4), 1718–1725.

Keating, P. A. (1990) The window model of coarticulation: articulatory evidence. In *Papers in laboratory phonology I*. (J. Kingston & M. E. Beckman, editors). Cambridge: Cambridge University Press, 451–470.

Keating, P. A. (1995) Segmental phonology and non-segmental phonetics. In *Proceedings of the XIIIth international congress of phonetic sciences*.

Kent, R. D. & Burkard, R. (1981) Changes in acoustic correlates of speech production. In *Aging: communication processes and disorders* (D. S. Beasley & G. A. Davis, editors), pp. 47–62. New York: Grune and Stratton.

Kent, R. D. & Moll, K. L. (1975) Articulatory timing in selected consonant sequences. *Brain & Language*, **2**, 304–323.

Kenyon, J. S. (1951) *American pronunciation*. Ann Arbor: Wahr.

Kerswill, P. E. (1985) A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers on Phonetics and Experimental Linguistics*, **4**.

Kingston, J. (1990) Articulatory binding. In *Papers in laboratory phonology I: between the grammar and physics of speech* (J. Kingston & M. E. Beckman, editors). Cambridge: Cambridge University Press, pp. 406–434.

Kohler, K. J. (1992) Gestural reorganization in connected speech: a functional viewpoint on "articulatory phonology." *Phonetica*, **49**, 205–211.

Krakow, R. A. (1989) *The articulatory organization of syllables: a kinematic analysis of labial and velar gestures*. Ph.D. dissertation, Yale University.

Kuehn, D. P. & Moll, K. L. (1976) A cineradiographic study of VC and CV articulatory velocities, *Journal of Phonetics*, **4**, 303–320.

Ladefoged, P. (1993) *A course in phonetics* (Third edition). New York: Harcourt, Brace, Jovanovich.

Ladefoged, P., DeClerk, J., Lindau, M. & Papçun, G. (1972) An auditory-motor theory of speech production, *UCLA Working Papers in Phonetics*, **22**, 48–75.

Lamontagne, G. A. (1993) *Syllabification and consonant cooccurrence conditions*. Ph.D. dissertation, University of Massachusetts.

Levine, H. (1960) Robust tests for equality of variance. In *Contributions to probability and statistics* (I. Olkin, editor), Palo Alto: Stanford University Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) Perception of the speech code, *Psychological Review*, **74**, 431–461.

Löfqvist, A. & Yoshioka, H. (1980a) Laryngeal activity in Swedish obstruent clusters, *Journal of the Acoustical Society of America*, **68**, 792–801.

Löfqvist, A. & Yoshioka, H. (1980b) Laryngeal-oral coordination in American English obstruent production, *Journal of the Acoustical Society of America*, **68**, S102(A).

Löfqvist, A. & Yoshioka, H. (1981a) Interarticulator programming in obstruent production, *Phonetica*, **38**, 21–34.

Löfqvist, A. & Yoshioka, H. (1981b) Laryngeal activity in Icelandic obstruent production, *Nordic Journal of Linguistics*, **4**, 1–18. (Also published in *Haskins Laboratories: Status Report on Speech Research*, **SR-63/64**, 1980, 275–292).

Löfstedt, I. (1992) Swedish segment length and structure preservation, *Studia Linguistica*, **46**(2), 93–127.

MacKay, I. R. A. (1978) *Introducing practical Phonetics*. Boston: Little, Brown.

MacNeilage, P. (1970) Motor control of serial ordering of speech, *Psychology Review*, **77**, 182–196.

Mattingly, I. G. (1981) Phonetic representation and speech synthesis by rule. In *The Cognitive Representation of Speech* (T. Myers, J. Laver & J. Anderson, editors). Amsterdam: North-Holland, 415–420.

McCarthy, J. & Prince, A. (1986) *Prosodic morphology*. ms., Department of Linguistics, University of Massachusetts, Amherst, and Program in Linguistics, Brandeis University, Waltham, Massachusetts.

Nolan, F. (1992) The descriptive role of segments: Evidence from assimilation. In *Papers in laboratory phonology II: gesture, segment, prosody* (G. J. Docherty & D. R. Ladd, editors). Cambridge: Cambridge University Press.

Öhman, S. E. (1967) Numerical model of coarticulation, *Journal of the Acoustical Society of America*, **41**, 310–320.

Petursson, M. (1977) Timing of glottal events in the production of aspiration after [s], *Journal of Phonetics*, **5**, 205–212.

Pierrehumbert, J. & Talkin, D. (1992) Lenition of /h/ and glottal stop. In *Papers in laboratory phonology II: gesture, segment, prosody* (G. J. Docherty & D. R. Ladd, editors), pp. 90–117. Cambridge: Cambridge University Press.

Prator, C. H. & Robinett, B. W. (1985) *Manual of American English pronunciation*. New York: Holt, Rinehart and Winston.

Prince, A. (1980) A metrical theory for Estonian quantity, *Linguistic Inquiry*, **3**, 511–562.

Ramig, L. (1983) Effects of physiological aging on speaking and reading rates, *Journal of Communication Disorders*, **16**, 217–226.

Recasens, D., Fontdevila, J. & Pallarès, M. D. (1993) An electropalatographic study of stop consonant clusters, *Speech Communication*, **12**, 335–355.

Ringel, R. L., Chodzko-Zajko & Offenbach, S. I. (1986) Age, health and the speech process. ASHA Mini-seminar, Detroit, Michigan.

Selkirk, E. O. (1982) The syllable. In *Structure of phonological representations*. (H. Van der Hulst & N. Smith, editors) *part 2*. Dordrecht: Foris Publications.

Steriade, D. (1982) *Greek prosodies and the nature of syllabification*. Ph.D. Dissertation, MIT.

Stevens, K. N. & House, A. S. (1963) Perturbation of vowel articulation by consonantal context: an acoustical study, *Journal of Speech and Hearing Research*, **6**, 111–128.

Stevens, K. N. & Keyser, S. J. (1989) Primary features and their enhancement in consonants, *Language*, **65**, 81–106.

Welford, A. T. (1977) Causes of slowing performance with age, *Interdisciplinary Topics in Gerontology*, **11**, 43–51.

Winer, B. J. (1971) *Statistical principles in experimental design*. New York: McGraw-Hill Book Company.

Winitz, H., Scheib, M. E. & Reeds, J. A. (1972) Identification of stops and vowels, *Journal of the Acoustical Society of America*, **51**(4), 1309–1317.

Yoshioka, H., Löfqvist, A. & Hirose, H. (1981) Laryngeal adjustments in the production of consonant clusters and geminates in American English, *Journal of the Acoustical Society of America*, **70**, 1615–1623.

Yoshioka, H., Löfqvist, A. & Hirose, H. (1982) Laryngeal adjustments in Japanese voiceless sound production, *Journal of Phonetics*, **10**, 1, 1–10.

Zec, D. (1989) *Sonority constraints on prosodic structure*. Ph.D. Dissertation, Stanford University.