ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/229439049

Phrasal signatures in articulation

Chapter · January 2000

CITATIONS
90

READS
22

4 authors, including:

Dani Byrd
University of Southern California

118 PUBLICATIONS **2,871** CITATIONS

SEE PROFILE

Shrikanth S Narayanan
University of Southern California

683 PUBLICATIONS **10,754** CITATIONS

SEE PROFILE

Elliot Saltzman
Boston University

129 PUBLICATIONS **4,539** CITATIONS

SEE PROFILE

Available from: Dani Byrd
Retrieved on: 13 May 2016

5

Phrasal signatures in articulation

DANI BYRD, ABIGAIL KAUN,
SHRIKANTH NARAYANAN,
AND ELLIOT SALTZMAN*

5.1 Introduction

As research in speech production becomes more integrated with linguistic theory, it has become increasingly clear that segmental articulation can not be understood independently of prosodic structure. We see evidence for prosodic structure in the physical act of articulation—that is, not just in what we say but in how we say it. These phonological influences pervade low-level articulatory behavior. Despite the pervasiveness of these effects, only a very few articulatory correlates of prosodic structure—what we call "prosodic signatures"—have been identified.

5.1.1 Phrasal structure and articulation

While speech scientists know a great deal about how individual segments are articulated—how they are realized in space and time—we know less about how words are put together in longer utterances. Just as position in the word and syllable affects the details of a gesture's articulation (Byrd, 1996a; Browman & Goldstein, 1995; Sproat & Fujimura, 1993; Krakow, 1989; Hardcastle, 1985), so does position in an utterance's phrasal structure. Phrasal structure affects the spatial and durational properties of individual articulatory gestures as well as (by extension) the temporal coordination of gestures.

With respect to the spatial domain, Fougeron and Keating (1996, 1997) find an increase in the magnitude of lingual gestures (as measured by linguapalatal contact) for consonants initial in increasingly large domains—word-initial < phonological (1997) or accentual (1996) phrase-initial < intonational phrase-initial or utterance-initial. Additionally, they find that articulations in domain-initial positions at each level are larger than medial and final positions at that level (Fourgeron & Keating, 1997). Dilley, Shattuck-Hufnagel, and Ostendorf (1996) present data on glottalization of word-initial vowels that demonstrates that phrase inital position is differentiated from phrase medial position by increased glottalization frequency, and find that intermediate and full intonational phrases are similarly differentiated. Pierrehumbert and Talkin (1992) report a larger glottal opening for [h] when it is initial after a phrase boundary than when it is in the middle of the phrase (but see also their comment on lengthening, p. 114).

The temporal patterning of an utterance is one way of encoding its meaning or its linguistic structure. Port, Cummins, and McAuley (1995) comment that the difference between saying "$2*(3^2)$" and "$(2 * 3)^2$" can be described in terms of the "location of valleys and peaks in the instantaneous speaking rate, brief decelerations or accelerations that lengthen or shorten speech segments...not usually a matter of pauses or silent gaps" (p. 342). One example of temporal patterning that has been well explored is phrase final lengthening—the phenomenon of longer acoustic or articulatory duration in the final position of some domain. Typically the last vowel before a large phrasal boundary is lengthened, but other units such as final consonants, VC's, syllables, and words have been identified as subject to domain-final lengthening as well. Acoustic data reported in Wightman, Shattuck-Hufnagel, Ostendorf, and Price (1992) indicate that several distinct types of phrasal boundaries can be distinguished by their degree of final lengthening. This result encourages the view that a variety of boundary strengths are active in determining speech timing. Byrd and Saltzman (1998) observe multiple levels of boundary-adjacent lengthening of articulatory gestures and model this lengthening in terms of gestural dynamics. Edwards, Beckman, and Fletcher (1991), Beckman, Edwards, and Fletcher (1992), and Beckman and Edwards (1992) report data on jaw movement that can be summarized as demonstrating a decreased gestural "stiffness," yielding increased duration and decreased peak velocity, of the final oral closing gesture of a VC sequence in phrase-final position.

Only a handful of studies have examined phrasal effects on patterns of relative timing among multiple gestures. Jun (1990) examined voice onset time (VOT), i.e., the temporal interval between a supraglottal event (release of a closure) and a laryngeal event (the onset of vocal fold vibration). She found that VOT increases from {word-medial, phrase-medial} position to {word-initial, phrase-medial} position to {word-initial, phrase-initial} position. She hypothesizes that it is the interaction of the spatial magnitude of the glottal gesture with the timing between gestures that yields the VOT differences. Similarly, longer VOTs are reported by Pierrehumbert and Talkin (1992) for stop consonants in phrase-initial versus phrase-medial position. Relatedly, McClean (1973) in his cinefluographic study finds that the onset of velum lowering in /CV#Vns/ sequences is consistently delayed with respect to the onset of lingual movement for the preceding tautosyllabic vowel in those cases where prosodically marked boundaries (i.e., major syntactic boundaries) intervene between the vowels, but not where word-only boundaries (e.g. subject#verb, adjective#noun) exist. Hardcastle (1985) reports electropalatographic data for /k#l/ sequences indicating that speaking rate and phrasing interact in affecting the timing of tongue body and tip raising. Generally, Hardcastle found that "the condition least favourable to co-articulation [between [k] & [l]] is the prosodically marked clause or sentence boundary at the ['normal slow' utterance rate]." These results imply that there is less temporal overlap between the /k/ and /l/ gestures at these prosodic boundaries. Holst and Nolan (1995) studied assimilation in [s]-[ʃ] sequences as inferred from an acoustic continuum that they categorize as ranging from most like a [sʃ] sequence, indicating an absence of assimilation, to a sequence with spectrally stable [ʃ] characteristics, indicative of assimilation. In her commentary on these data, Browman (1995) concludes that the degree of gestural overlap between the consonants was negatively correlated with the presence of an intervening clause boundary. In sum, these studies suggest that phrasal position is a significant force in constraining the degree of temporal coproduction among articulatory gestures.

Thus, it is the case that the temporal and spatial characteristics of articulatory gestures are governed in part by their phrasal position. This conception of position-dependency is fundamentally a linguistic notion, and an examination of articulatory detail has much to gain by synthesizing linguistic concepts with detailed study of speech kinematics and dynamics. The data presented in Section 5.2 will consider the temporal organization and spatial detail of oral articulatory gestures in the immediate, or local, neighborhood of a phrase boundary. In turn we will interpret these prosodic effects on articulation using the task dynamics model of gestural control developed by Saltzman and Munhall (1989).

5.1.2 A dynamical systems model of articulation

In the task dynamics model of speech production (Saltzman & Munhall, 1989), a damped mass-spring equation of motion is hypothesized to control articulatory gestures. Such an equation of motion is given in (i):

$$m\ddot{x} + b\dot{x} + k\left(x - x_{\text{targ}}\right) = 0\,, \qquad (i)$$

where x, $\dot{x}$, and $\ddot{x}$ are position, velocity, and acceleration, respectively; m is the mass parameter (generally assumed to be unit mass), b is the damping coefficient, k is the spring stiffness,[1] and x_{targ} is the target position. Since the model generally assumes critical damping, the following form of equation (i) is useful, in which ξ appears as the damping ratio (equal to 1 for critical damping) and ω_0 is the undamped natural frequency:

$$\ddot{x} + 2\xi\omega_0\dot{x} + \omega_0^2\left(x - x_{targ}\right) = 0\,,$$

$$\text{where } \omega_0 = \sqrt{k/m} \;\&\; x = b/2m\omega_0 = b/2\sqrt{km}. \qquad \textit{(ii)}$$

The relation of these terms to the spatiotemporal properties of the movements that they model is discussed further in Section 5.4.

In this framework, each gesture is associated with an activation interval (most simply instantiated as a step function, but see Byrd & Saltzman, 1998) whose strength defines the degree to which the gesture shapes the vocal tract at any given point in time. The relative timing of two gestures is a result of the temporal coordination of their activation intervals. Gestural coproduction occurs when the activation intervals of two or more gestures overlap in time. If one gesture's activation period is prematurely ended due to coproduction with a following gesture, the first gesture is said to be truncated by the second (see for example Bullock & Grossberg 1988, Nittrouer et al. 1988, Harrington et al. 1995).[1]

5.2 Method

The articulatory phenomena investigated below include boundary-adjacent lengthening, inter-articulator relative timing across a boundary, and magnitude differences due to phrasal position. Among the questions addressed by this experiment are the following. First, does position at the edge of a phrasal domain affect the temporal and spatial characteristics of individual oral consonantal gestures? Based on the findings outlined above, we expect consonant gestures to lengthen in the neighborhood of phrase boundaries. And second, is the degree of temporal coproduction between gestures affected by phrasal structure? Again based on the findings outlined above, we expect coproduction of consonant gestures to decrease when the gestures span a phrasal boundary.

The experimental subject (SN) was a speaker from Madras of the Brahmin dialect of Tamil. Movement tracking with a magnetometer was used to examine the articulatory kinematics of two nasal consonant sequences spanning a word boundary, [n#m] and [m#n], in a variety of Tamil sentences.[2] The boundary at the juncture between the consonants was manipulated such that three conditions were included: a word boundary (possessor-possessed or compound noun), a small phrase boundary (subject#object)[3], and a large phrase boundary (vocative name followed by a request). These experimental conditions will be referred to as WORD, SMALL PHRASE, and LARGE PHRASE respectively. The 23 sentences included are shown (using orthographic conventions favored by the subject) in Appendix A. WORD and SMALL PHRASE sentences consist of a single intonational phrase (IP), whereas LARGE PHRASE sentences contain two. The intonational contour of these sentences consists of a high pitch accent on the syllable bearing phrasal stress followed by a low boundary tone that associates to the right IP edge, giving rise to an overall pattern of falling pitch within an IP.

Fifteen repetitions of the sentences in a pseudo-random order, yielding 345 analyzed tokens, were recorded using the EMMA magnetometer system for transducing the movement of small coils attached to the articulators (Perkell, Cohen, Svirsky, Matthies, Garabieta & Jackson, 1992; see also Gracco & Nye, 1993, Löfqvist, 1993). Transducers were placed on the nose, maxilla, jaw, upper and lower lips (at the vermilion border), tongue tip, and on three locations on the tongue body. For this study, only vertical (y) position signals for the tongue tip and upper lip are considered.[4] (NB: For the remainder of this paper we will speak of movement of an articulator (e.g. tongue tip movement) which should be understood as referring to movement in the midsagittal plane of the transducer placed on that articulator.) The movement data were sampled at 625 Hz with low-pass filtering at 300 Hz before voltage-to-distance conversion. After voltage to distance conversion (with a filter cutoff of 17 Hz), correction for head movement (using the nose and maxillary reference transducers), and rotation to the occlusal plane, the position signals underwent 25 point smoothing by a triangular filter. The upper lip and tongue tip vertical position signals were differentiated to yield velocity signals, also smoothed at 25 points.

The *HADES* signal analysis program (Rubin, 1995) was used to identify algorithmically the times of onset, extremum, and end of vertical movement for each consonant in the cluster, that is, upper lip for [m] and tongue tip for [n]. These events were defined by identifying the zero-crossings of the velocity signals. The time and articulator vertical-position at each of these points were recorded. (The onset and end points were defined to be not at absolute zero crossing but at the point at which velocity reached 10% of the maximum closing (for onsets) or opening (for ends) velocity for that articulator across the entire set of utterances. This criterion was established in order to obtain “crisp” zero-crossings as velocity sometimes hovered near zero. (see Saltzman, Löfqvist, Kinsella-Shaw, & Rubin, 1992.) If more than one velocity maximum or minima occurred during a closing (or opening) movement of an articulator, that closing or opening gesture was excluded from analysis (8 movements out of 1380 were excluded for this reason). (One token was excluded due to subject error.) Additionally, for each opening and closing movement, the magnitude and timepoint of the peak velocity was recorded. Finally, the acoustic duration of the nasal sequence was measured from the waveform. Criteria for acoustic segmentation included changes in amplitude (as indicated in an accompanying energy plot) and the apparent absence of higher resonant frequencies in the spectrum. The presence or absence of a voicing break during the nasal sequence was also noted, as was the length of any such break.

The data analysis considers the effect of boundary condition on C1 closing, C1 opening, and C2 closing (where C1≡preboundary consonant & C2≡postboundary consonant).[5] Additionally, the temporal coordination of the C1 and C2 closing movements are examined. Based on the time and magnitude values for movement onset, extremum, and end, a variety of dependent variables were calculated. These variables include:

a) the durations of closing (onset to extremum) and opening (extremum to offset) movements for each consonant;
b) the time between onsets (Δonsets) & extrema (Δextrema) of closing movements;
c) the displacements (onset to extremum & extremum to offset) for each consonant;
d) the time from the onset of an opening or a closing movement to its peak velocity.

Measures a–c are shown in Figure 5.1. Additionally, two relative timing measures (i.e., measures normalized for consonant closing duration), shown in Figure 5.1 (right), were calculated—C2-onset-inside-C1 (the percent of the way into C1 closing that C2 onset occurred) and C1-extremum-inside-C2 (the percent of the way into C2 closing that C1 displacement extremum occurred). For example, for the measure C1-inside-C2, a smaller number means that C1

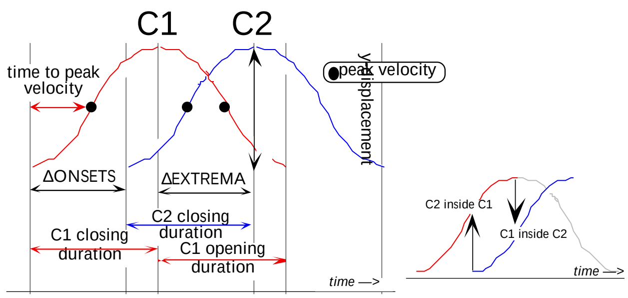

Figure 5.1. *A schema of the vertical movement of the upper lip (inverted) and tongue tip articulations showing the measurements calculated.*

reached its extremum position proportionally earlier in C2, indicating less coproduction. As a whole, these variables reflect the duration and magnitude of the individual consonantal articulations and the intergestural timing of the consonants. In conjunction with the peak velocity information, these data can help inform us as to the dynamics that underlie the consonant sequence articulation.

Two-factor analysis of variance tests the effects of boundary condition (3-levels) and sequence (2-levels) on the dependent measure. When the main effect of boundary is significant ($p \leq .05$), planned comparisons of means test for significant differences between each pair of boundary conditions for each sequence with a confidence criterion of $p \leq .05$. The hypotheses tested in the planned comparisons are that durations will lengthen and consonant coproduction decrease from the WORD to the SMALL PHRASE to the LARGE PHRASE condition.

5.3.0 Results

5.3.1 Acoustic Data

The acoustic duration of the nasal sequence was found to be significantly longer in the LARGE PHRASE condition than in both other conditions for both sequences ($F(2,331)=55.061$, $p<.0001$), with a mean duration of 122 ms, as compared to 94 ms and 93 ms for WORD and SMALL PHRASE respectively. Voicing was generally continuous through the nasal sequence, with only 10 of 345 tokens having a voicing break; all of these in the LARGE PHRASE condition. This fact and our own listening confirm that substantial pauses are rare in the data set.

5.3.2 Boundary-Adjacent Lengthening

There is only a marginal main effect of boundary condition on the duration of the closing movement for the preboundary consonant, C1 ($F(2,336)<3$, $p=.052$). (The planned comparisons of means demonstrated that for the [m#n] sequence the LARGE PHRASE boundary condition had slightly longer (≈ 6ms) C1 closing duration.) Next, we consider lengthening of the opening of the first consonant, which we view as associated with the initiation of the second phrase. There is a significant effect of boundary on C1 opening duration ($F(2,334)=50.96$, $p<.0001$). C1 opening duration is significantly longer in the LARGE PHRASE boundary condition than in the SMALL PHRASE and WORD boundary conditions for both sequences. Finally, the duration of the postboundary consonant's (C2's) closing gesture is significantly affected by the type of preceding boundary ($F(2,338)=42.482$, $p<.0001$) such that the LARGE PHRASE boundary condition is significantly longer that the other two conditions for both consonant sequences. There is also a significant interaction of boundary and sequence ($F(2,338)=4.601$, $p=.0107$) due to the fact that the difference was more robust for the [m#n] sequence.

5.3.3 Spatial Magnitude

Only small effects on gestural magnitude were observed. No effect of boundary type is found for C1 closing displacement. Boundary has a significant effect on C1 opening displacement ($F(2,334)=62.398$, $p=.0001$). There is also a significant interaction effect ($F(2,334)=50.58$, $p=0001$). In the [n#m] sequence, the C1 opening displacement in the LARGE PHRASE condition was significantly greater than in both other conditions, and WORD was also greater than SMALL PHRASE. However, the differences in displacement are very small; at or near the approximately 0.5mm spatial resolution of EMMA. Finally, the displacement of the domain-initial, that is postboundary (C2) consonant closing movement is significantly affected by boundary ($F(2,337)=10.063$, $p=.0001$). There is also a significant interaction of sequence and boundary ($F(2,337)=3.971$, $p=.0197$). Planned comparisons determine that for [n] (i.e., the [m#n] sequence) the LARGE PHRASE condition displacement is bigger than the other two conditions, although the differences are again quite small in magnitude.

5.3.4 Intergestural Timing

5.3.4.1 Absolute Timing

Next consider the temporal organization of the domain final and domain initial consonants, specifically: the time between onsets (Δonsets) and the time between extrema (Δextrema). (See Fig. 1) There is a small effect of boundary on Δonsets ($F(2,336)=4.603$, $p=.0107$). The planned comparisons demonstrate that the LARGE PHRASE condition has a slightly longer (≈ 7ms) Δonsets than the other two conditions for the [n#m] sequence, and a marginally ($p=.0838$) longer (≈ 6ms) Δonsets than the SMALL PHRASE condition for the [m#n] sequence. Thus, it seems that phrasal category has only a very small effect on how the onset of C1 and C2 are coordinated temporally. By contrast, boundary condition has a large effect on the time between the extrema of the consonants ($F(2,336)=32.467$, $p<.0001$) such that Δextrema is significantly longer (≈ 145ms) in the LARGE PHRASE boundary condition than in the SMALL PHRASE and WORD boundary conditions for both consonant sequences. The Δextrema means are shown in Figure 5.2 (left).

Lastly, one might expect that as a sequence spans a greater number of prosodic domain edges there would be greater variability in its intergestural timing; that is, that the temporal cohesion between gestures would decrease. In this experiment the timing between the consonants is more variable when they span a (large) phrase boundary. The standard deviation for Δextrema in the LARGE PHRASE condition is about twice those of the other conditions. This suggests that there is less constraint on the intergestural timing of consonants when they are in separate phrasal domains.

5.3.4.2 Relative Timing

C2-inside-C1 and C1-inside-C2 index the relative timing between the consonant closures. Based on the results above—that is, minimal change in C1 closing duration and in Δonsets—we expect only a small effect on C2-inside-C1 but a stronger effect on C1-inside-C2 comparable to that on C2 closing duration. Accordingly, we find a small effect of boundary on C2-inside-C1 for only the [n#m] sequence ($F(2,336)=3.346$, $p=.0364$) such that the LARGE PHRASE condition is less overlapped than the other two conditions (a difference of approximately 7%). Boundary has a stronger effect on C1-inside-C2 ($F(2,336)=13.156$, $p<.0001$) such that C1 peaks proportionally earlier in the C2 closure. That is, C1 is less overlapped with C2 in the LARGE PHRASE boundary condition than in the SMALL PHRASE and WORD boundary conditions for the [n#m] sequence (a difference of about 10%). In the [m#n] sequence, LARGE PHRASE is less overlapped with SMALL PHRASE (a difference of about 5%).

5.3.5 Time to Peak Velocity

Movement peak velocity information can help illuminate the dynamics underlying observed durational patterns. Specifically, we consider the time from the movement onset to its peak velocity in order to determine if the observed lengthening patterns are consistent with variation in the gestural stiffness parameter.[6]. This measure is relevant because in a mass-spring gestural model, if the gestural stiffness differs in the two conditions, the time to peak velocity should differ. Gestures with lower stiffness will have later occurring peak velocities than ones with higher stiffness.

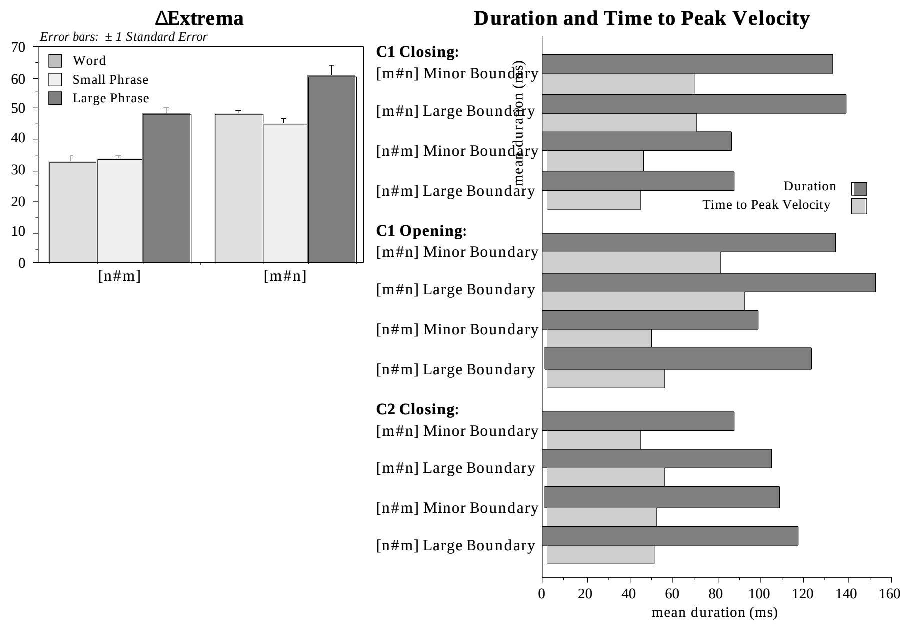

Figure 5.2 LEFT: ΔExtrema split by sequence and phrasal condition; RIGHT: The relation between duration and time to peak velocity.

In this analysis, we pool together the small phrase and word boundary conditions because duration in the large phrase condition differed from the other conditions for each movement, while the word and small phrase condition did not differ.[7] No significant effect of boundary condition is found for the C1 closing movement. Significant differences for C1 opening ($F(1,336)=52.39$ $p<.0001$) and C2 closing ($F(1,340)=20.932$, $p<.0001$) exist, with the LARGE BOUNDARY condition having a longer time to peak velocity. A significant interaction effect of boundary and sequence on C2 closing time to peak velocity ($p<.0001$) indicates that this result is characteristic only of the [m#n] sequence. The relation between duration and time to peak velocity is shown in Figure 5.2 (right).

5.3.6 Summary of Results

To summarize, we observed boundary-adjacent lengthening of the opening of the preboundary consonant and the closing for the postboundary consonant with concomitant small increases in gestural magnitude and less temporal coproduction between final and initial consonant articulations belonging to separate phrasal domains. These findings are interpreted in Section 4 by considering the behavior in the LARGE PHRASE boundary versus MINOR BOUNDARY groups. These results are summarized in Table 5.1.

Table 5.1. *Duration, displacement, and time to peak velocity results (MB=minor boundary, LP=large phrase).*

movement	consonant	duration	displacement	time to peak velocity
C1 closing		main effect	no effect	no effect
[n#m]	n	--	--	--
[m#n]	m	LP>MB	--	--
C1 opening		main effect	main effect	main effect
[n#m]	n	LP>MB	LP>MB	LP>MB
[m#n]	m	LP>MB	--	LP>MB
C2 closing		main effect	main effect	main effect
[m#n]	n	LP>MB	LP>MB	LP>MB
[n#m]	m	LP>MB	--	--

5.4.0 Dynamical Bases of Lengthening

Our interests lie ultimately in understanding the dynamical underpinnings of linguistically-conditioned duration changes. In this regard, the prosodically motivated durational differences observed in these data could arise from a number of differences in the underlying dynamics. Using the task dynamics model of articulatory gestures (Saltzman & Munhall, 1989; see Section 5.1.2 above), we find that independent manipulation of the model’s gestural parameters (see Equation [ii])—natural frequency (stiffness), target, damping ratio—has the kinematic consequences shown in Table 5.2. Some of these effects are outlined graphically in Beckman and Edwards (1992) and Beckman, et al. (1992). Target, stiffness, and truncation differences are the primary types of parameter changes we entertain here as mechanisms for lengthening. Table 5.3 also includes the consequences of gestural truncation and of "linear rescaling" (Harrington et al., 1995).

A change in a gesture's target position alone (or, equivalently, a change in initial displacement from a given target) will result in a difference in gestural displacement and magnitude of peak velocity, but will yield no differences either in gestural duration or in the time from gestural initiation to peak velocity. These kinematic properties are simply consequences of the critically-damped, mass-spring dynamics hypothesized to underlie gestural control (see Equation [ii]).

Table 5.2. *Summary of kinematic consequences of various mass-spring equation parameter manipulations.*

mass-spring parameters	measured kinematic variables				
	duration	**displacement** (max × min)	**time to peak velocity**	**peak velocity**	**avg velocity** (disp/time)
natural frequency					
less (lower ω_0)	longer	no change	longer	lower	lower
more (higher ω_0)	shorter	no change	shorter	higher	higher
target					
bigger	no change	greater	no change	higher	higher
smaller	no change	smaller	no change	lower	lower
truncation					
less	longer	greater*	no change	no change	likely to decrease
more	shorter	smaller*	no change	no change	likely to increase
damping ratio ($0<\xi<1$)					
more damping	longer	less overshoot	longer	lower	lower
less damping	shorter	more overshoot	shorter	higher	higher
target & natural frequency scaled proportionally (c.f. "linear rescaling" (Harrington et al., 1995) where amplitude & duration are scaled proportionally)					
less shrinking	longer	greater	longer	no change	no change
more shrinking	shorter	smaller	shorter	no change	no change

(*These changes may be small if the gesture has a plateau-like shape at its displacement extremum.)

Within this model, the dynamical parameter that most directly controls movement duration is the stiffness of a gesture. The stiffer a gesture is (i.e., the larger the value of k in Equation [i]), the faster its associated articulators move for a given initial displacement from its target. Assuming that a gesture's activation remains on at least through the point at which the gesture has reached its target, a gesture with lower stiffness will take longer to reach the target than a gesture with higher stiffness, regardless of the initial displacement from target. Additionally, lower stiffness gestures will display longer times from gestural initiation to peak velocity than higher stiffness gestures. As with the effects of varying gestural target position (or initial displacement from target), these kinematics result from the damped mass-spring dynamics used to model gestural control.

The relative timing of two gestures is a result of the temporal coordination of their activation intervals. The truncation of one gesture due to (a canonically following) but overlapping gesture will cause the first gesture to terminate before it reaches its target and, hence, display a shorter gestural duration with no change in time to peak velocity (excepting severely truncated cases). The change in displacement may be small, however, if the gesture has a plateau-like shape at its displacement extremum.

There are other means by which durational variations can be generated using a damped mass-spring gestural model. We merely note a few of these here. First, changes in a gesture's damping ratio will give rise to variations in the gesture's duration. However, the task dynamic model as implemented by Browman and Goldstein (1990) and others generally assumes the dynamics are invariantly specified to be critically damped. As noted by Beckman and Edwards (1992), work by Smith, Browman, McGowan, & Kay (1993) generally supports this assumption of critical damping across a variety of speech gestures. A second possible manipulation is "linear rescaling" outlined by Harrington et al. (1995). This refers to what is, in effect, a uniform spatial and temporal scaling of gestural kinematics. Scaling that contracts a gesture spatially and temporally will result in a shorter duration, smaller displacement, and shorter time to peak velocity; but peak velocity itself (as well as average velocity) will be unaffected. Lastly, kinematic durational changes may result from variations in the time course of gestural activation. Various models implement gestural parameter values not as step functions but as functions with gradually changing onsets and offsets, e.g. a ramped function used to define trajectories of a gesture's stiffness or target (e.g. Ostry, Gribble, & Gracco, 1996; Byrd & Saltzman, 1998; see also Kröger, Schröder, & Opgen-Rhein, 1995 for a related treatment). Such changes in the course of parameter-value trajectories can give rise to differences in resultant movement durations.

One possibility that we consider further in Section 5.5.1 is that prosodic structure influences ongoing articulatory activity by inducing temporally local modulations of gestural parameter values. If, for example, prosodic influences induced a local drop in gestural stiffness after the gesture attained its peak velocity, the time to peak velocity would be unaffected, but the kinematics would be altered after that point.

5.4.1 Lengthening of C1 opening and C2 closing

A comparison of the predicted kinematic changes in Table 5.2 with the observed kinematic changes in Table 5.1 suggests that the lengthening of the C1 opening and

C2 closing and C1 opening (for [m]) adjacent to a large phrase boundary appear to be largely due to local lowerings of gestural natural frequency at the large phrase boundary. This is indicated by the "signature" change in time to peak velocity. This may be accompanied by some very small increases in target (or, equivalently, decreases in truncation) for [n]. The lengthening of C2 closing for [m] is more difficult to interpret, as time to peak velocity did not change. Truncation cannot be ruled out here, despite the fact that displacement did not change.

To summarize, we have observed the following kinematic phenomena in considering nasal consonant sequences spanning a large phrase boundary:

(1) Boundary-adjacent lengthening for the opening of the preboundary consonant and for the closing for the postboundary consonant; and

(2) Less temporal coproduction between word-final and -initial consonant articulations belonging to distinct phrasal domains.

We have interpreted the kinematic data in terms of the parameter manipulation in a mass-spring gestural model. In most cases, the changes in temporal patterning of the articulations due to their phrasal position appear to be due to local lowerings of gestural stiffness at the edges of high-level prosodic domains.

5.5 From Phonology to Dynamics

The fact that we see evidence of phrasal structure in the spatiotemporal patterning of articulation means that, if we take seriously the goal of adopting the task dynamics model to language, we must seek an explicit understanding of the connection between prosodic hierarchical structure and inter- and intra- gestural dynamics. We consider next the role of prosodic structure in determining gestural parameter values. (For a discussion of the role of linguistic structure in the assignment of intergestural coordination, see Byrd, 1996b.)

If we acknowledge that gestural parameters such as target, stiffness, and (more abstractly) activation interval differ as a function of the prosodic position of a gesture, then we must take seriously Keating's (1995) concern that "lexical specification [cannot tell] us how to pronounce a word, only how to pronounce it in some particular context" (p. 31). We must explain how these parameter values can be determined in a principled way as a function of post-lexical structure. Moreover, as noted in Fougeron and Keating (1996), since the influence of prosodic position has been shown to apply to multiple articulatory subsystems—including tongue tip, lips, jaw, and velopharyngeal opening—prosodic information must be a high-level component of speech motor control. The details of this incorporation of prosodic structure into gestural dynamics is likely to be language and speaker specific (e.g.

Byrd & Saltzman, 1998). We have argued that the articulators become infused with the prosodically appropriate behavior by modulation of parameter values associated with the lexically specified constriction goals (i.e., gestures) that are achieved by the articulators. This captures the local changes in speaking rate that we observed at phrase edges.

For the purpose of addressing phonological and morphological problems, speech is organized into discrete phonemes whose internal structure is atemporal. These phonemic units are organized into a variety of hierarchical structures such as syllables (and/or subsyllabic units) and phrases (e.g. PP and IP) that are active in linguistic processes. In contrast, for the purpose of studying speech as complex and communicative human movement, speech is organized into goal-directed units of action—gestures. These gestures have intrinsic durational characteristics and are temporally coordinated in an overlapping manner. Given this state of affairs, the finding that post-lexical structure and articulation interact leads inevitably to the following question: *How can inherently atemporal symbolic units exert, in a principled fashion, the effects need to explain the observed patterning of low-level, inherently temporal, action units?*

5.5.1 A dynamical implementation of phrasal structure

We suggest that in addition to being seen as symbolic and atemporal, phrasal boundaries can also be seen as displaying inherent duration, that is, as having a temporal domain over which they exert their influence on parameter values of the active articulatory gestures. This approach is a first step in conceiving a dynamical implementation of phrasal structure.

Our primary concern here is edge effects, that is, changes in articulatory gestural dynamics that occur at the edges of prosodic domains. Characterizing edge effects requires a means of computing prosodic boundary strength at a juncture. Various methods for such computation have been proposed in the literature. In the SPE model (Chomsky & Halle, 1968), for instance, boundary symbols are inserted between various syntactic categories, and boundary strength is reflected in the number of boundary symbols. (See also Cooper & Paccia-Cooper, 1980.) In more recent work (e.g. Selkirk, 1978, 1984 and Nespor & Vogel, 1982, 1986), however, a separate representational level is posited, referred to typically as p-structure, that relates syntactic constituency to phonological structure. In p-structure, prosodic constituents are organized hierarchically, constituting domains for the application of phonological rules..[8] Within such a framework "boundary strength" can be computed by identifying the highest level constituent-edge at a given juncture.

We hope to wed the theoretic notion of prosodic structure with a dynamical model of speech production since it is these dynamics that we suggest are influenced by prosodic context. Specifically, in the most current formulation of the task dynamics model of speech production (see Saltzman, 1995, Rubin et al., 1996, Saltzman, this volume), a recurrent, sequential network architecture (Jordan, 1990; see also Lathroum, 1989) is adopted to pattern gestural activation trajectories over time. Additionally, an explicit "clocking" subnetwork that controls ongoing speech rate is incorporated into the architecture of the overall system (Saltzman, this volume). This subnetwork acts to modulate the system's "clock rate," that is, the amount of absolute time defined between each successive "tick" of internal network computation. The resultant time-scaling is equivalent to modulating the stiffnesses of all currently active gestures (Saltzman, this volume).

How might prosodic context serve to modulate speech dynamics within such a model? Saltzman (1995) suggests that syllable position effects on the organization of multi-gestures segments such as [l] and [n] might be accounted for by a non-tract-variable (i.e. non-constriction-based) boundary element. We expand on this concept by proposing that such a unit occurs at prosodic domain edges and has its activity governed by prosodic constituency. We refer to such prosodic boundary units as *π-gestures*. The activation level of a π-gesture is specified as a function of prosodic boundary strength. Saltzman (1995) suggests that activation levels of tract-variable gestures may be affected in proportion to the strength (i.e. activation level) of the π-gesture. We suggest here that a given π-gesture independently and directly (i.e., not mediated via tract-variable gesture activations) affects: (i) the values of gestural parameters such as stiffness or target position for all tract-variable gestures with which it is concurrently active; or (ii) the clock rate (i.e., local speaking rate) such that a stronger π-gesture yields more slowing of the clock rate than a weaker one.

The initial, conservative hypothesis is that only one type of π-gesture exists for domain edges. This predicts that different levels in the prosodic hierarchy should not be accompanied by edge effects different in kind. That is, different prosodic boundaries will be realized with *qualitatively* identical dynamic consequences that differ only in *degree* as a function of the activation level of the π-gesture. Moreover, there is no notion of "left" versus "right" edge; that is, "left" and "right" are not primitives of the model. This further predicts that only the temporal organization of the π-gesture with respect to the constriction-based gestures can potentially yield different initial versus final edge effects. Differen.tial amounts of anticipatory or carryover coproduction of the π-gesture with overlapping constriction-based gestures would yield differential (and presumably language-specific) amounts of domain-initial versus domain-final changes in the gestures' parameters (e.g. stiffness).

Finally, the dynamic implementation of phrasal structure predicts that only the gestures within the π-gesture's temporal field of activation would be directly affected, not gestures remote from the phrasal boundary. This, intuitively, seems in accordance with the quite local nature of domain edge effects that have been observed. An understanding of the factors that shape the duration of π-gestures and their precise coordination with tract-variable gestures remains work for future investigation.

5.6 Conclusion

This presentation has examined the spatial and temporal patterning of oral articulatory gestures as a function of phrasal structure for two nasal consonant sequences in Tamil. This and other work demonstrates that prosodic structure is manifest in the details of articulation. The precise nature of phrasal effects on both intergestural timing and gestural duration and magnitude remains an open question; as does the cross-linguistic typology of and constraints on these perturbations. We have argued for the necessity of integrating the abstract symbolic representation useful to linguists with a dynamical model of human movement useful to speech scientists. The ultimate goal of this research effort is to determine general "signatures" of prosodic structure on articulatory organization and capture the empirical reality of these signatures through the modulation of a small number of parameters of the speech production system.

Appendix

WORD	[n#m]	Gugan magan Santosh.	Santosh is Gugan's son.
		Gugan maga Shanti.	Shanti is Gugan's daughter.
		Murugan magaranna peru.	It's called Murugan Peak.
	[m#n]	Arangam nagam kuurpu.	Arangam's fingernail is sharp.
		Arangam nargaril Santosh irukan.	Santosh lives in Arangam town.
		Kanagam nagam kuurpu.	Kanagam's fingernail is sharp.
		Kanagam nargaril Santosh irukan.	Santosh lives in Kanagam town.
SMALL PHRASE	[n#m]	Gugan magesha pathan.	Gugan saw Magesh.
		Gugan magendrana pathan.	Gugan saw Magendran.
		Murugan magesha pathan.	Murugan saw Magesh.
		Murugan magendrana pathan.	Murgan saw Magendran.
	[m#n]	Arangam nakirana pathan.	Arangam saw Nakiran.
		Arangam nagulana pathan.	Arangam saw Nagulan.
		Kanagam nakirana patha.	Kanagam saw Nakiran.
		Kanagam nagulana patha.	Kanagam saw Nagulan.
LARGE	[n#m]	Gugan, magana inga kutindu va.	Gugan, bring your son here.

PHRASE	Gugan, magadathuku polam.	Gugan, let's go to Magada.
	Murugan, magana inga kutindu va.	Murugan, bring your son here.
	Murugan, magadathuku polam.	Murugan, let's go to Magada.
[m#n]	Arangam, nagaru angendu.	Arangam, move away from there.
	Arangam, nagatha kadikathe.	Arangam, don't bite your nails.
	Kanagam, nagaru angendu.	Kanagam, move away from there.
	Kanagam, nagatha kadikathe.	Kanagam, don't bite your nails.

References

Basri, H., Broselow, E., Finer, D., & Selkirk, E. (1998) "Affix classes in Makassarese: Exploiting the word/stem distinction" a talk presented by E. Selkirk at Yale University, May 28, 1998.

Beckman, M.E. & J. Edwards. 1992. Intonational categories and the articulatory control of duration. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (pp. 359-375). Tokyo: Ohmsha.

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

Browman, C.P. 1995. Assimilation as gestural overlap: comments on Holst and Nolan. In B. Connell & A. Arvaniti (Eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. (pp. 334-342). Cambridge: Cambridge University Press.

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

Browman, C.P. & L. Goldstein. 1995. Gestural syllable position effects in American English. In F. Bell-Berti and L. Raphael (Eds.), *Producing speech: Contemporary issues. For Katherine Safford Harris* (pp. 19-34). Woodbury, NY: AIP Press.

Bullock, D. & S. Grossberg. 1988. Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review 95*: 49-90.

Byrd, D. 1996a. Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24(2):209-244.

Byrd, D. 1996b. A phase window framework for articulatory timing. *Phonology 13(2)*:139-169.

Byrd, D. & E. Saltzman. 1998. Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics 26:173-199*.

Chomsky, N. & M. Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Cooper, W. and J. Paccia-Cooper. 1980. *Syntax and Speech*. Cambridge: Harvard University Press.

Dilley, L., S. Shattuck-Hufnagel, & M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*, 423-444.

Docherty, G. J. 1992. *The timing of British English obstruents*. Berlin; New York: Foris Publ., (Netherlands phonetic archives; 9).

Edwards, J., M.E. Beckman, & J. Fletcher. 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society, 89*, 369-382.

Fougeron, C. & P. Keating. 1996. Variations in velic and lingual articulation depending on prosodic position: Results for 2 French speakers. *UCLA Working Papers in Phonetics, 92*, 88-96.

Fougeron, C. & P. Keating. 1997. Articulatory strengthening in prosodic domain-initial position. *JASA 101(6)*:3728-3740.

Gracco, V.L. & P.W. Nye. 1993. Magnetometry in speech articulation research: Some misadventures on the road to enlightenment. *Forschungberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München 31*: 91-104.

Hardcastle, W.J. 1985. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication, 4*, 247-63.

Harrington, J., J. Fletcher, & Roberts. 1995. Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics 23*: 305-322.

Holst, T. & F. Nolan. 1995. The influence of syntactic structure on [s] and [ʃ] assimilation. in B. Connell and A. Arvanti (Eds.) *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. (pp. 315-333) Cambridge: Cambridge University Press.

Jordan, M.I. 1990. Motor learning and the degrees of freedom problem. In M. Jeannerod, (Ed.). *Attention and Performance XIII*. Hillsdale, NJ: Erlbaum.

Jun, S-A. 1995. Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell and A. Arvanti (Eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. (pp. 235-253) Cambridge: Cambridge University Press.

Keating, P.A. 1995. Segmental phonology and non-segmental phonetics. In K Elenius & P. Branderud (eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences* 3. 26-32.

Krakow, R.A. 1989. *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures*. Ph.D. dissertation, Yale University.

Kröger, B., G. Schröder, & C. Opgen-Rhein. 1995. A gesture-based dynamic model describing articulatory movement data. *Journal of the Acoustical Society of America 98.4* 1878-1889.

Lathroum, A. 1989. Feature encoding by neural nets. *Phonology, 6*, 305-316.

Löfqvist, A. 1993. Electromagnetic transduction techniques in the study of speech motor control. *PHONUM: Reports from the Department of Phonetics, University of Umeå 2*: 87-106.

McClean, M. 1973. Forward coarticulation of velar movement at marked junctural boundaries. *J. Speech Hear. Res., 16*(2), 286-296.

Nespor, M. & I. Vogel. 1982. Prosodic domains of external sandhi rules. in H. van der Hulst and N. Smith (eds.) *The Structure of Phonological Representations*. vol. 1, Dordrecht: Foris, pp. 222-255.

Nespor, M. & I. Vogel. 1986. *Prosodic Phonology*. Dodrecht: Foris Publications.

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

Ostry, D. J., P. Gribble, & V.L. Gracco. 1996. Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned? *The Journal of Neuroscience*, *16(4)*, 1570-1579.

Perkell, J., M. Cohen, M. Svirsky, M. Matthies, I. Garabieta, & M. Jackson. 1992. Electromagnetic midsagital articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, *92*, 3078-3096.

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

Port, R. F., F. Cummins, & J.D. McAuley. 1995. Naive time, temporal patterns, and human audition. In T. van Gelder and R. Port (Eds.), *Mind as Motion*. (pp. 339-371) Cambridge: MIT Press.

Rubin, P.E. 1995. HADES: A case study of the development of a signal analysis system. *Applied speech technology*. Boca Raton, FL, CRC Press. 501-520.

Rubin, P.E., E. Saltzman, R. McGowan, L. Goldstein, M. Tiede, & C. Browman. 1996. *CASY and extensions to the task-dynamics model.* in the Proceedings of the First ESCA Tutorial and Research Workshop on Speech Production Modeling, Autrans, France.

Saltzman, E. 1995. Intergestural timing in speech production: Data and modeling. In K. Elenius and P. Branderud (eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences 2*. 84-91.

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

Saltzman, E., A. Löfqvist, J. Kinsella-Shaw, & P.E. Rubin. 1992. A perturbation study of lip-larynx coordination. *Proceedings of the International Conference on Spoken Language Processing (ICSLP '92): Addendum*. Edmonton, Canada, Priority Printing. 19-22.

Saltzman, E., A. Löfqvist, & S. Mitra. this volume. “Clocks” and “glue”—global timing and intergestural cohesion. Papers in Laboratory Phonology V.

Selkirk, E.O. 1978. On prosodic structure and its relation to syntactic structures. *Nordic Prosody*, 2, 111-140.

Selkirk, E.O. 1984. Phonology and Syntax: The Relation between Sound and Structure. (MIT, Cambridge, MA).

Smith, C.L., C.P. Browman, R.S. McGowan, & B. Kay. 1993. Extracting dynamic parameters from speech movement data. *Journal of the Acoustical Society of America 93*: 1580-1588.

Sproat, R. & O. Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, *21*, 291-311.

Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf, & P. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91* 1707-1717.

Notes

[*]The authors gratefully acknowledge the assistance and advice of two anonymous reviewers and the editors, Louis Goldstein, Vince Gracco, Walter Naito, and the support of NIH Grants HD-01994, DC-00121, & DC-00016 to Haskins Laboratories. And thanks to C. Wiltshire for inviting us to collect pilot data for this project.

i In accordance with Browman & Goldstein (1987; see also Kröger, Schröder, & Opgen-Rhein 1995), we hypothesize that the initiation and termination of each gesture's activation interval are specified according to the evolving phase angles of a virtual undamped mass-spring oscillator with the same natural frequency (ω_0) as the critically damped gestural system. Each gesture's activation interval begins at a phase angle of 0^{o} and ends at a phase angle for which the gesture is within a criterion distance from its target. Thus, changes in gestural stiffness will be associated with corresponding changes in the virtual oscillator's ω_0, with the consequence that lower stiffness gestures will have intrinsically longer activation intervals (due to lower ω_0s) than higher stiffness gestures.

The relative timing of two gestures is specified according to specific phase angles (or phase windows, Byrd 1996b; see also Docherty 1992) of each gesture's virtual oscillatory cycle. For example, two gestures might be phased relative to each other so that 90^{o} in the second gesture's virtual cycle is synchronized with 270^{o} in the first gesture's virtual cycle. An increase in the second gesture's phase angle, or a decrease in the first gesture's phase angle, results in an increased amount of temporal overlap or coproduction between the gestures.

i In the subject's dialect final [m] is sometimes realized as nasalization and rounding on the preceding vowel. In the vast majority of the tokens analyzed here, final [m] is realized with an actual closure. In all cases lip adduction occurs.

i Basic word order in Tamil is SOV; thus the object initiates a maximal projection, i.e., the verb phrase.

i Movement of the upper lip, as opposed to lower lip or lip aperature, was selected for analysis in order to prevent the introduction of jaw movement as a confounding factor.

i The C2 opening showed no consistent effects of the types discussed below. This may be due either to an actual lack of systematic boundary effects on C2 opening or, more trivially, to the fact that the postboundary words vary more in segmental content later in the word (i.e., after the word initial [#Cag...]).

i An additional measure that has been used to evaluate potential changes in gestural stiffness is the slope of the regression line fit to data points in the peak velocity vs. displacement plot. Different slopes have been interpreted as indicating gestures of different stiffnesses. However, this measure is valid as an indicator of stiffness only when a gesture in fact reaches its target, i.e., when there is no truncation. Furthermore, this relation is linear with a zero-intercept only for gestures with no damping. For this reason, we choose to evaluate time to peak velocity.
i A two-factor ANOVA testing for effects of boundary (now 2-level: large boundary vs. minor boundary) and of sequence on movement duration confirms the (3-level) ANOVA results reported above.
i Current work suggests that the requirements for exhaustivity and non-recursivity in prosodic parsing may be ranked as violable constraints (Basri, Broselow, Finer, & Selkirk, 1998).

Notes

[*]The authors gratefully acknowledge the assistance and advice of two anonymous reviewers and the editors, Louis Goldstein, Vince Gracco, Walter Naito, and the support of NIH Grants HD-01994, DC-00121, & DC-00016 to Haskins Laboratories. And thanks to C. Wiltshire for inviting us to collect pilot data for this project.

1 In accordance with Browman & Goldstein (1987; see also Kröger, Schröder, & Opgen-Rhein 1995), we hypothesize that the initiation and termination of each gesture's activation interval are specified according to the evolving phase angles of a virtual undamped mass-spring oscillator with the same natural frequency (ω_0) as the critically damped gestural system. Each gesture's activation interval begins at a phase angle of 0^o and ends at a phase angle for which the gesture is within a criterion distance from its target. Thus, changes in gestural stiffness will be associated with corresponding changes in the virtual oscillator's ω_0, with the consequence that lower stiffness gestures will have intrinsically longer activation intervals (due to lower ω_0s) than higher stiffness gestures.

The relative timing of two gestures is specified according to specific phase angles (or phase windows, Byrd 1996b; see also Docherty 1992) of each gesture's virtual oscillatory cycle. For example, two gestures might be phased relative to each other so that 90^o in the second gesture's virtual cycle is synchronized with 270^o in the first gesture's virtual cycle. An increase in the second gesture's phase angle, or a decrease in the first gesture's phase angle, results in an increased amount of temporal overlap or coproduction between the gestures.

2 In the subject's dialect final [m] is sometimes realized as nasalization and rounding on the preceding vowel. In the vast majority of the tokens analyzed here, final [m] is realized with an actual closure. In all cases lip adduction occurs.

3 Basic word order in Tamil is SOV; thus the object initiates a maximal projection, i.e., the verb phrase.

4 Movement of the upper lip, as opposed to lower lip or lip aperature, was selected for analysis in order to prevent the introduction of jaw movement as a confounding factor.

5 The C2 opening showed no consistent effects of the types discussed below. This may be due either to an actual lack of systematic boundary effects on C2 opening or, more trivially, to the fact that the postboundary words vary more in segmental content later in the word (i.e., after the word initial [#Cag...]).

6 An additional measure that has been used to evaluate potential changes in gestural stiffness is the slope of the regression line fit to data points in the peak velocity vs. displacement plot. Different slopes have been interpreted as indicating gestures of different stiffnesses. However, this measure is valid as an indicator of stiffness only when a gesture in fact reaches its target, i.e., when there is no truncation. Furthermore, this relation is linear with a zero-intercept only for gestures with no damping. For this reason, we choose to evaluate time to peak velocity.

7 A two-factor ANOVA testing for effects of boundary (now 2-level: large boundary vs. minor boundary) and of sequence on movement duration confirms the (3-level) ANOVA results reported above.

8 Current work suggests that the requirements for exhaustivity and non-recursivity in prosodic parsing may be ranked as violable constraints (Basri, Broselow, Finer, & Selkirk, 1998).