Cho, Taehong (2006). In L. Goldstein, D. H. Whalen, C. T. Best (Eds.), Laboratory Phonology 8, pp.519–548. Mouton de Gruyter

Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English

Taehong Cho

This study investigates effects of prosodic structure on kinematic variations that may illuminate how prosody is manifested in articulatory variation. Kinematic characteristics of lip aperture (e.g., articulatory displacement, duration and velocity) were examined with respect to three prosodically strong locations: domain-initial, and domain-final syllables. The results (obtained with the Electromagnetic Articulograph, EMA) showed that each of these prosodic locations are associated with distinctive kinematic patterns that can distinguish itself from others: All of the three prosodically important locations showed strengthening effects with generally longer and larger movements in common, but they differred primarily in velocity: faster for accented gestures, no change for domain-initial/final lip opening gestures, and slower for cross-boundary lip closing gestures. This suggests that a hierarchically-nested prosodic structure is marked by systematic kinematic variation. The results were further evaluated in terms of the dynamical parameter settings (e.g., stiffness, intergestural timing, target amplitude, re-scaling) in a mass-spring gestural model. Close examination of relationships between various kinematic parameters (e.g., displacement, velocity, duration, time-to-peak velocity) suggested that at the very least one should look for a combination of settings for multiple dynamical parameters, in order to account for the prosodically-driven articulatory systematicity in a way that is both descriptively and explanatorily adequate.

1. Introduction

The term *prosody* has traditionally been used as the cover term for suprasegmental features such as pitch, duration and intensity, but more recently, it is often used to refer to more abstract linguistic notion “a hierarchically organized structure of phonologically defined constituents and heads" (Beckman 1996), in which lower domains (e.g., syllables) are typically grouped into immediately higher levels (e.g., words), eventually forming the Intonational Phrase (IP) (see Shattuck-Hufnagel and Turk 1996 for a review). This structural view of prosody assumes that prosody is a grammatical entity in its own right, which is a crucial element of speech production and speech comprehension processes.

One line of research, taking the structural view of prosody, has vigorously shown that the prosodic structure of an utterance is phonetically manifested on the surface by distinctive pitch patterns and temporal structure at the right edge of prosodic constituent (e.g., Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; Edwards, Beckman and Fletcher 1991; Jun 1993; 1998, and Beckman 1996). Further, a recent articulatory study has shown that domain-final vocalic articulation in English also undergoes a substantial increase in spatial magnitude (Cho 2002, 2005), contrary to a common assumption that domain-final position is subject primarily to temporal expansion (cf. Beckman, et al. 1992).

While these domain-final phonetic events have widely been considered as major phonetic correlates of prosodic structure, researchers have recently started to look at domain-initial positions for other potential phonetic events correlated with prosodic structure (Pierrehumbert and Talkin 1992; Jun 1993; Fougeron and Keating 1997; Cho and Keating 2001; Fougeron 2001; Cho 2002, 2004, 2005; Keating, Cho, Fougeron and Hsu 2003). For example, a phrase-initial stop /t/ is likely realized with a longer VOT and a larger linguopalatal contact than the same /t/ occurring in the middle of a phrase, a phenomenon know as *domain-initial strengthening*.

Yet other phonetic correlates of prosodic structure other than domain-final phenomena come from cross-boundary phenomena such as cross-boundary vowel-to-vowel coarticulatory resistance (Cho 2004) and the relative timing of consonant and vowel gestures (Byrd 2000). For example, in an articulatory study, Cho (2004) showed that vowels in prosodically stronger locations are coarticulated less with neighboring vowels, but do not exert a stronger influence on the articulation of neighboring vowels.

Such robust phonetic phenomena in the vicinity of prosodic boundaries have led to a growing awareness that it is no longer fruitful to describe the sound properties of a language without adequately taking into account the interface between prosodic structure and phonetics. Accordingly, the focus of recent laboratory work has been on more diverse prosodic locations, including domain-initial and -final positions, as well as stressed (pitch-accented) syllables (de Jong 1991, 1995, Cho 2002). These three positions have been shown to give rise to some type of strengthening of articulatory properties of features or gestures (also known as *prosodic strengthening*), which is taken to be an articulatory signature of prosodic structure.

The majority of phonetic research has, however, generally been limited to one prosodic effect (the stress or the edge effect), and has thus failed to provide a comprehensive account of articulatory characteristics of prosodic structure. The present study examines the three prosodic locations concurrently, in order to understand the prosody-phonetics interface in a coherent way. In particular, it aims to understand the effects of prosody in English on kinematic variations and considers dynamical accounts that may illuminate how prosody is manifested in articulatory variation. To this end, it examines lip movement kinematics of accent- and boundary-induced articulatory strengthening and how accent-induced kinematic patterns differ from boundary-induced ones. Further, given that prosodically-conditioned articulatory variation may be controlled by a particular dynamical parameter setting in a mass-spring gestural model (Beckman, et al. 1992; Harrington, et al. 1995; Byrd and Saltzman 1998; Byrd, et al. 2000), it is further evaluated whether and how prosodically-driven strengthening may be accounted for by a particular dynamical parameter setting, and whether different dynamical mechanisms govern the articulatory characteristics that arise from different prosodic locations.

Task dynamic model and dynamical parameters. In the task dynamic model, the articulatory gesture is described in terms of the behavior of an abstract 'mass' (an idealized articulator such as the tongue) which is connected to a 'spring' and a 'damper' in a critically damped mass-spring system (Saltzman and Munhall 1989). As Hawkins (1992) describes, it is as if one end of the spring is attached to the mass, and the other end is held at the target location. Then, as the target location changes, the spring is stretched, and the mass is pulled towards the target location. In a critically damped mass-spring gestural model, however, the mass does not oscillate, but asymptotes towards the equilibrium position, such that the gesture is generally realized as a one-directional movement towards the target.

In the model, the gesture is defined as a dynamical system specified with a set of parameter values. Relevant dynamical parameters include target (underlying amplitude), stiffness (or natural frequency), damping ratio, intergestural timing, and activation time. Characteristics of the articulatory movements that result from executing gestures depend on the values of the parameters specified for a given gesture. Crucially, any systematic kinematic variation is interpreted as the consequence of dynamical parameter settings. Thus, in theory, systematic kinematic variations arising from prosodic conditions should be accounted for by either a particular dynamical mechanism or an interaction of more than one mechanisms.

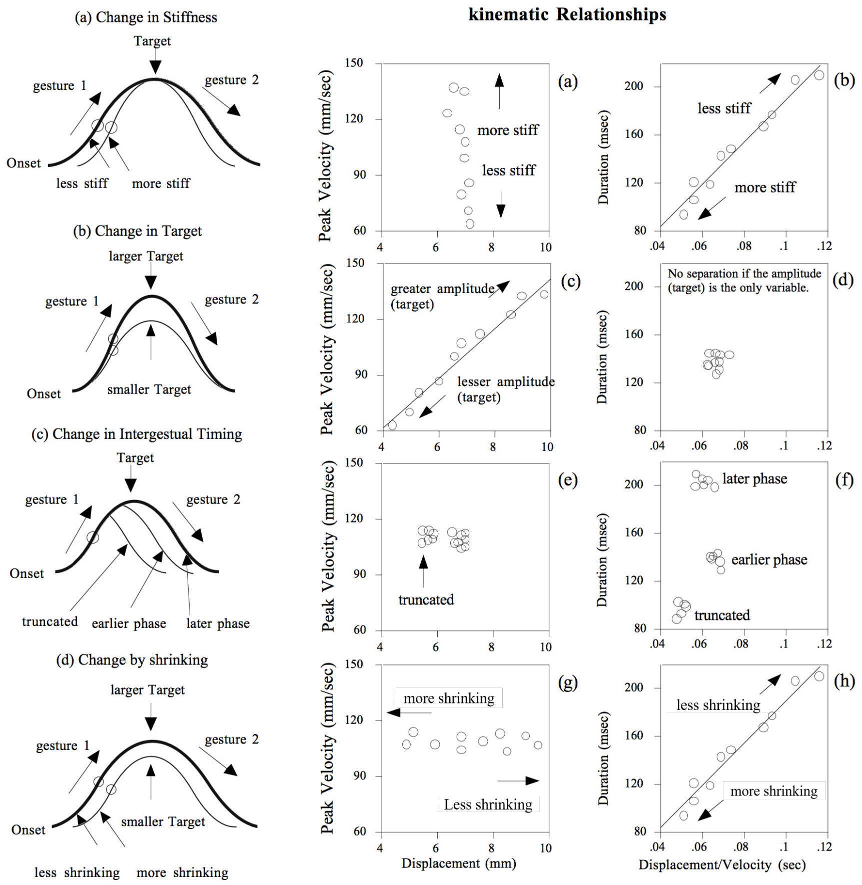

Figure 1. Hypothetical trajectories that correspond to a change in each parameter (left) and relationships among kinematic variables that manifest dynamical parameter settings (right). In the left panel, empty circles indicate the timepoint of the peak velocity attainment. In the right panels, (a–b) show change in stiffness; (c–d), change in amplitude (target); (e–f), change in intergestural timing; and (g–h), change by shrinking. Figures (a)–(f) are adapted from Beckman, et al. (1992).

Some researchers (Beckman, et al. 1992; Byrd, et al. 2000) provide useful summaries of the kinematic consequences of various mass-spring parameter manipulations. The left panel of Fig. 1 shows schematized movement trajectories that correspond to changes in four dynamical parameters. (Although shrinking is a scaling of two parameters, it will be called 'parameter' for the sake of simplicity.) The right panels of Fig. 1 visualize idealized kinematic manifestations of different dynamical specifications by relating some kinematic measures to each other.

(1) *Stiffness.* Variation in movement duration is thought to be controlled by the stiffness parameter: the stiffer the spring (the articulator), the faster the movement (see the left panel of Fig. 1). If stiffness is the only parameter underlying kinematic differences, there should be a change in peak velocity (the maximum velocity during the movement), but not in displacement (spatial distance that the articulator travels), therefore showing vertical distribution of the datapoints (Fig. 1a). In addition, there should be a proportional change in both duration and displacement/velocity ratio, with a diagonal distribution of the datapoints (Fig. 1b), i.e., as duration increases, peak velocity would decrease, making the displacement/velocity ratio increase (because of constant displacement). Further, the time-to-peak-velocity (the interval from the onset to the attainment of peak velocity) will vary as stiffness changes (the less stiff, the longer) (cf. Byrd and Saltzman 1998; Byrd, et al. 2000; and Byrd 2000).

(2) *Target (articulatory amplitude).* A change in target induces a change in displacement. In a pure target change, peak velocity and displacement changes proportionally without a change in duration: With stiffness being held constant, articulators have to travel farther with no extra time. The only way to reach the increased target is by increased velocity in proportion to the change in the target value, with a diagonal distribution of the datapoints (Fig. 1c). Further, since displacement and velocity change proportionally, there should be no change in the displacement/velocity ratios, nor should there be a change in duration (due to no stiffness change) (Fig. 1d). Time-to-peak-velocity also remains constant (the left panel of Fig. 1).

(3) *Intergestural timing or truncation.* The articulatory movement towards the target can be 'truncated' by an early activation of the following gesture, which keeps the movement from reaching its assumed target. Thus, under a pure change in intergestural timing, there should be no change in peak velocity because the effect of a substantially earlier following gesture is mainly to prevent the preceding gesture from reaching its target. As Byrd, et al. (2000) noted, if the gesture has a plateau-like shape at its peak displace-

ment and the truncation applies primarily to this region, the change in displacement will be small or zero, but relatively larger if the truncation applies to the region beyond the plateau-like shape (See Fig. 1e). Further, there will be a substantial change in duration as the following gesture is phased earlier or later, whereas the ratio of the displacement to the peak velocity remains relatively unchanged because of no change in displacement and velocity, except when enough truncation brings about a decrease in displacement (Fig. 1f). Finally, the durational change comes from a change in the interval from the timepoint of peak-velocity to the target (deceleration duration) with no change in time-to-peak-velocity.

(4) *Shrinking.* Shrinking can be defined as a change in both target and stiffness which are scaled proportionally. Shrinking can be thought of as a unique dynamical parameter that may underlie prosodically conditioned kinematic variation (see Harrington, et al. (1995), and Byrd, et al. (2000)). As can be seen in Fig. 1 (left), in a pure proportional change in target and stiffness, there would be a proportional increase in both duration and displacement, which results in no change in peak velocity, giving a horizontal distribution of the datapoints (Fig. 1g). Further, the displacement/velocity ratio will increase as duration increases, giving a diagonal distribution of the datapoints (Fig. 1h). Note that the pattern in Fig. 1h is similar to that in Fig. 1b under a change in stiffness. However, the difference between these two is that in a change in stiffness (Fig. 1b), the displacement/velocity ratio increases as duration increases not because displacement increases, but because velocity decreases with displacement being held constant.

2. Experiment

In order to examine effects of various prosodic conditions on speech production, lip movement data in American English were collected, using Electromagnetic Midsagittal Articulography (Carstens Articulograph). An important criterion for building the corpus was to include both prosodic and segmental variables. Each item in the corpus included two test syllables (pre- and post-boundary), yielding a $C_1V_1\#C_2V_2$ sequence (# = a prosodic boundary) across words. C_1 and C_2 were always /b/, whose articulation is known to minimally interfere with the vocalic lingual articulation. V_1 and V_2 were either /i/ or /a/, resulting in four pairs: /bi#bi/, /ba#ba/, /bi#ba/, and /ba#bi/, but in this study only the data for /bi#bi/ and /ba#ba/ are examined to control for the vowel type.

The boundary between the test syllables was varied from the Intonational Phrase boundary (IP), to the Intermediate Phrase boundary (ip), to the Word boundary (Wd). At the same time, accentuation was manipulated in preboundary and postboundary syllables, resulting in four patterns: ACC#ACC, ACC#UNACC, UNACC#ACC, UNACC#UNACC. Such a manipulation yields three prosodic variables: (a) prosodic boundary; (b) accentuation of syllables (accented, unaccented); (c) position of test syllables (initial, final). Thus, the corpus contained every combination of the prosodic and segmental factors, yielding a total of 24 different sequences (3 boundaries 2 accentual patterns in the preboundary syllable 2 accentual patterns in the postboundary syllable 2 vowel type (/bibi/ vs. /baba/). (Sample sentences are given below in Table 1.)

Six American English speakers participated. In order to control for rounding in the low vowel, speakers whose dialect lacked /ɔ/ were chosen. They were all trained in producing English sentences in the ToBI framework (Beckman and Elam 1997) prior to the experiment. Before the actual recording date, each speaker participated in an approximately two-hour long practice session in order to be able to produce the intended renditions as naturally as possible. During the experiment, speakers were instructed to produce two different versions (ip vs. IP) of a sentence in order to obtain balanced ip and IP tokens.

The target sequences were obtained from real sentences in a mini discourse situation intended to induce the desired variety of accent-placement patterns and prosodic groupings. A sample sentence set with /ba#ba/ tokens in *Little Bah bopped the girl* is given in Table 1. (/bi#bi/ sequence tokens were produced in similar discourse frames as in *Donna B. beeped at him*) In each target sentence, the words in bold received pitch accent. The prompt was read silently by the speaker to cue the intended accent patterns, which were provided using partial ToBI transcriptions in the script. Six American English speakers were recorded reading the target sentences. Each sentence was read twice in succession, and the entire list was read twice, for a total of four repetitions per sentence. This yielded a total of 576 sentence tokens (24 sentence types x 6 speakers x 4 repetitions).

In the EMA experiment, seven transducer coils were used. Two reference transducers were placed on the nose and upper gumline, or maxilla, in order to correct for head movement inside the helmet. Two transducers were mounted on the upper and lower lips at the vermilion borders (L1, L2) to monitor lip closing and opening movements. (The remaining three transducers were located on the tongue; the data from these transducers were analyzed in Cho (2002, 2004, 2005)).

Table 1. A subset of the corpus containing /ba#ba/ sequences with different prosodic boundaries (IP, ip, Wd) and accentual patterns.

\# = Word boundary:

(a) *ACC.-UNACC.*	
Prompt:	Did you just say "Little **Boo** bopped the girl last night"?
Target:	No, "Little **Bah** # bopped the girl"
rendition:	H* L– L%
(b) *UNACC.-ACC.*	
Prompt:	Did you just say "Little Bah **popped** the girl last night"?
Target:	No, "Little Bah # **bopped** the girl"
rendition:	H* L– L%
(c) *ACC.-ACC.*	
Target:	You know what? Little **Bah** # **bopped** the girl.
rendition:	H* H* L– L%
(d) *UNACC.-UNACC.*	
Prompt:	Did you just say "**Big** Bah bopped the girl last night"?
Target:	No, "**Little** Bah # bopped the girl"
rendition:	H* L– L%

\# = Intermediate or Intonational Phrase boundaries (ip or IP):

(e) *ACC.-UNACC.*	
Prompt:	Did you say "Little **Boo** bopped the **boy** last night"?
Target:	No, " Little **Bah** # bopped the **girl**."
rendition 1:	H*L– H* L– L%
rendition 2:	H*L–L% H* L– L%
(f) *UNACC.-ACC.*	
Prompt:	Did you say "**Big** Bah **popped** the girl last night"?
Target:	No, " **Little** Bah # **bopped** the girl.
rendition 1:	H* L– H* L– L%
rendition 2:	H* L–L% H* L– L%
(g) *ACC.- ACC.*	
Prompt:	Did you say "Little **Boo popped** the girl"?
Target:	No, " Little **Bah** # **bopped** the girl.
rendition 1:	H*L– H* L– L%
rendition 2:	H*L–L% H* L– L%
(h) *UNACC.-UNACC.*	
Prompt:	Did you say "**Big** Bah bopped the **boy** last night"?
Target:	No, " **Little** Bah # bopped the **girl**"
rendition 1:	H* L– H* L– L%
rendition 2:	H* L–L% H* L– L%

Next, the articulatory space was rotated so that the x-axis was the occlusal plane using a bite-plate with two additional transducers on it. The EMA data were sampled at 500 Hz and were then submitted to low-pass filtering with a filter cutoff of 50 Hz.

The relevant C_1V_1#C_2V_2 portion of the audio recording was transcribed, with the aid of an acoustic display, by two trained ToBI transcribers (one the author) following the criteria in the ToBI transcription system (Beckman and Elam 1997). The two transcribers identified the same locations of pitch accent in every token of the entire dataset. The only difference between the transcribers came from a choice between IP and ip. Because the difference is an important experimental variable in this study, only tokens whose renditions were agreed on by the two transcribers were used. (There was a 96.3% agreement between the two transcribers in distinguishing ip and IP boundaries.)

2.1. Measurements

To obtain lip opening and closing movement data, horizontal and vertical position signals for two lip sensor coils are combined into one dimension, Lip Aperture. The Euclidean distance between these two sensor coils is used as an index of Lip Aperture (cf. Byrd and Saltzman, 1998). The derived signal serves as the basis for all the lip measurements. The onset and target timepoints of the lip closing and opening movements were determined from the zero-crossings in the velocity signal with a velocity noise window, defined as 5% of the highest peak velocities of each lip opening and closing movements across the entire dataset. This procedure was done separately for each of the six speakers.

Various dependent variables are calculated at/between the moments of movement onset, target, and peak velocity. The measured variables that are examined in this paper are schematized in Fig. 2. There are three movement events (opening-closing-opening) and five different measures are made for each lip opening and closing movement: (a) *displacement (mm):* the spatial difference between the onset and the target (C_1-to-V_1 lip opening displacement; V_1-to-C_2 lip closing displacement; and C_2-to-V_2 lip opening displacement); (b) *total movement duration*: the temporal interval from the onset to the target (C_1ONS-TO-V_1TARG; V_1ONS-TO-C_2TARG; and C_2ONS-TO-V_2TARG); (c) *time-to-peak-velocity* (acceleration duration): the temporal interval from the onset to the timepoint of peak velocity (C_1ONS-TO-V_2PKVEL; V_1ONS-TO-

C_2PKVEL; and C_2ONS-TO-V_2PKVEL); (d) *deceleration duration*: the interval from the timepoint of peak velocity to the target (C_1PKVEL-TO-V_2TARG; V_2PKVEL-TO-C_2TARG; and C_2PKVEL-TO-V_2TARG); (e) *peak velocity*: the actual peak velocity value for each opening and closing lip movement.

Based on these measured variables, the relationships between some of them were examined further, in order to investigate detailed dynamical aspect of prosodic effects as discussed above (see Fig. 1).

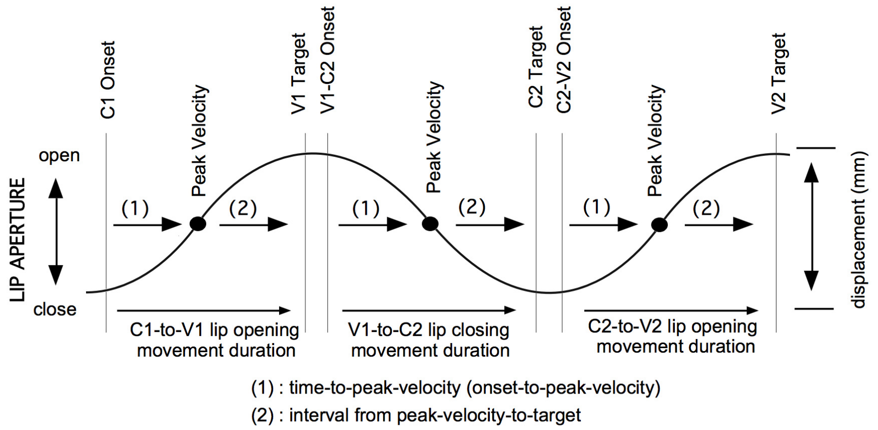

Figure 2. Schema of the lip opening and closing movement trajectory with an indication of the measured kinematic variables.

The systematic influence of prosodic factors on lip opening and closing gestures were evaluated, based on repeated measures Analyses of Variance (ANOVAs). The within-subject factors are V_1 Accent (ACC, UNACC), V_2 Accent (ACC, UNACC), and Boundary Type (IP, ip, Wd). The results are reported based on three-way ANOVAs performed separately for /a/ and /i/. To avoid violating the sphericity assumption (for Boundary Type with more than two levels), a Huynh-Feldt corrected degree of freedom was used in generating F ratio and p values. Next, for relationships between kinematic variables, simple regression analyses were performed. Since we are interested in overall patterns across speakers, and since each speaker had different magnitude of absolute measurements, data were normalized across speakers by transforming measured kinematic values into percentages. This returns for each datapoint that datapoint's percentage contribution to the sum of the entire dataset, which makes the variables more comparable across speakers.

3. Results

This section reports on the effects of Accent and Boundary Type on the kinematics of (1) a preboundary (domain-final) C_1-to-V_1 lip opening gesture, (2) a postboundary (domain-initial) C_2-to-V_2 lip opening gesture, and (3) a cross-boundary V_1-to-C_2 lip closing gesture. It should be noted that in this paper, only overall effects across speakers are reported. In general, for statistically significant findings reported in this study, speakers showed similar patterns. Due to the space limit, this paper will focus on the main effects. (For further details, readers are invited to refer to Cho (2002)).

3.1. Accent effects on kinematics

Let's first look at the results of lip *opening* movement. The basic finding is that lip opening movements in accented CV sequences are characterized by an increase, relative to unaccented sequences, in all measured variables (i.e., displacement, total movement duration, time-to-peak-velocity, deceleration duration, peak velocity), as summarized in Table 2. That is, when accented, lip opening movements are simply bigger in all ways regardless of position-in-domain (final vs. initial) and vowel type (/bi/ vs. /ba/).

Table 2. Summary of effects of accent on kinematics for domain-final C_1-to-V_1 and domain-initial C_2-to-V_2 lip opening gestures. The description in each cell (e.g., larger, longer, higher) is based on main effects, showing a pattern for the accented CV vs. the unaccented counterpart.

Kinematic measures	domain-final (C_1-to-V_1#)		domain-initial (#C_2-to-V_2)	
	/ba#/	/bi#/	/#ba/	/#bi/
Displacement	larger F=19.98**	larger F=88.18**	larger F=29.541**	larger F=41.82**
Total Movement Duration	longer F=37.41**	longer F=74.75**	longer F=79.40**	longer F=60.77**
C_1ons-To-V_1pkvel (time-to-peak-vel)	longer F=44.18**	longer F=103.8**	longer F=51.61**	longer F=61.27**
V_1pkvel-To-V_1targ (deceleration)	longer F=16.32**	longer F=38.07**	longer F=54.86**	longer F=43.98**
Peak Velocity	higher F=15.92**	higher F=62.58**	higher F=17.81**	higher F=37.69**

(** $p<0.01$, degrees of freedom = F[1,5])

Turning to lip *closing* movement, as shown in Table 3, V_1-to-C_2 lip closing gestures are influenced by both V_1 and V_2 Accent factors. With respect to the V_1 Accent effect, the results show patterns similar to those of lip opening gestures: V_1-to-C_2 lip closing gestures are associated with an increase in displacement, duration, time-to-peak-velocity, and peak velocity (showing a larger, longer, and faster movement).

As for the V_2 Accent effect, it also influences V_1-to-C_2 lip closing gestures with respect to several kinematic measures, but in a way that is somewhat different from V_1 Accent. First, V_1-to-C_2 lip closing gestures before accented V_2 are larger, but not faster, showing an increase in displacement for both /a#b/ and /i#b/, but this time with no change in time-to-peak-velocity. Second, there is a significant change in deceleration duration as a function of V_2 Accent, which was not the case for the effect of V_1 Accent. Finally, the total movement duration is not consistently longer when V_2 is accented: Only /a#b/ (not /i#b/) shows an increase in duration together with an increase in peak velocity.

Table 3. Summary of effects of V_1 and V_2 accent on kinematics for V_1-to-C_2 lip closing gesture. The description in each cell (e.g., larger, longer, higher) is based on significant main effects, showing a pattern for the accented CV as compared with the unaccented counterpart.

	When V_1 accented		When V_2 accented	
	/a#b/	/i#b/	/a#b/	/i#b/
Displacement	larger F=8.944*	larger F=37.882**	larger F=28.251**	larger F=8.691*
Total Movement Duration	longer F=6.354*	longer F=20.384**	longer F=6.414*	*n.s.*
V_1ONS-To-C_2PKVEL	longer F=6.567*	longer F=19.172**	*n.s.*	*n.s.*
VC_2PKVEL-To-C_2TARG	*n.s.*	*n.s.*	longer F=32.477**	longer F=11.833*
Peak Velocity	higher F=4.924*	higher F=8.923*	higher F=11.431*	*n.s.*

(** p<0.01, * p<0.05, ^{tr} p<0.07; degrees of freedom = F[1,5])

Dynamical aspects of Accent. One of the underlying assumptions in a task dynamics model is that distinct kinematic patterns that might arise from linguistic factors can be characterized by different settings of a specific dynamical parameter. Under this assumption, there arises a question as to what dynamical parameter can best characterize accent-induced kinematic variation. Some investigators (Edwards, et al. 1991; Beckman, et al. 1992; Harrington, et al. 1995) have already suggested that accent-induced kinematic variation in jaw opening movements is best captured by a single dynamical parameter, *intergestural timing*. However, we found no evidence that this intergestural timing account or any other parametric account can be extended to the accent-induced kinematic patterns.

Let's first consider lip opening movements. Regarding ACC/UNACC differences in lip opening gestures, we found that an accented lip opening gesture is associated with an increase in all kinematic parameter values (the longer, larger, faster pattern). When these results are compared to kinematic consequences of various mass-spring parameter manipulations, there seems to be no single specific mass-spring parameter that can account for ACC/UNACC differences (compare with the predictions in Fig. 1): (a) If intergestural timing were the only dynamical parameter, we would have observed an increase in both displacement and duration but no change in time-to-peak-velocity and peak velocity; (b) If gestural target (or underlying amplitude) were the only dynamical parameter, there would have been no change in total movement duration and time-to-peak-velocity; (c) In a pure change in stiffness, there would have been no change in displacement but a decrease in peak velocity for accented gestures; (d) Finally, in a pure change by shrinking, there would have been no change in peak velocity. However, none of these idealized descriptions matches the results presented here. Furthermore, relationships between various kinematic variables revealed that no particular dynamical parameter setting can be singled out as an absolute dynamical mechanism underling ACC/UNACC kinematic differences. (Due to the space limit, regression plots are not shown here. For a full description of the data, please see Cho, 2002.)

Next, for the cross-boundary V_1-to-C_2 *lip closing* gesture, kinematic patterns are different depending on the source of Accent (preboundary vs. postboundary) and Vowel Type. On the one hand, the effect of preboundary (V_1) accent shows a pattern similar to the effect of accent on the lip opening gesture, favoring no dynamical account. On the other hand, the fact that only the second component of the total duration (i.e., deceleration duration) is influenced by V_2 Accent appears to support the intergestural timing account, which is especially true for /i#b/ with no change in peak velocity and time-to-peak-velocity. (Recall that the patterning of no change in peak velocity along with an increase in displacement fits the descriptions of a delayed intergestural timing (see Fig. 1)). However, this intergestural timing account is critically weakened for /a#b/ which shows a change in peak velocity (the larger, the faster). The kinematic relationships also show that there is substantial overlapping between ACC and UNACC datapoints, not matching any idealized pictures for a pure change in any particular dynamical parameter. (Again, figures are not provided.)

3.2. Boundary effects on kinematics

C_1-to-V_1# (domain-final) lip opening gesture. The pattern of kinematics in common to both /ba#/ and /bi#/ is that a C_1-to-V_1 lip opening gesture before a higher boundary is associated with an increase in total movement duration, time-to-peak-velocity and deceleration duration, with no increase in peak velocity, i.e., showing a longer, but neither faster nor slower movement. Statistical results are summarized in Table 4. Furthermore, although there is no main effect of Boundary on displacement for /ba#/ (Table 4), there is a significant Accent x Boundary interaction (F[1.3,6.4]=5.99, p<0.05) because of a pattern of IP>(ip=Wd) only when /ba#/ is *unaccented* (Bonferroni/Dunn *posthoc* test). That is, the C_1-to-V_1 lip opening gesture is generally larger (increased displacement) before a higher boundary for both /ba#/ and /bi#/ except when /ba#/ is accented.

#C_2-to-V_2 (domain-initial) lip opening gesture. As in the case of C_1-to-V_1 (domain-final) lip opening gesture, a C_2-to-V_2 lip opening gesture after a higher boundary is associated with an increase in total movement duration, time-to-peak-velocity (C_1ONS-To-V_1PKVEL) with no increase in peak velocity, again showing a longer, but neither faster nor slower movement. This time, however, there is no effect of Boundary on deceleration duration, suggesting that the temporal effect lies primarily in the first component of the duration (i.e., time-to-peak-velocity).

With respect to displacement, there is no main effect of Boundary. However, there is a significant Boundary x Accent interaction for both /#ba/ and /#bi/ (F[1.6,8.1]=6.345, p<0.025; F[2,10]=4.65, p<0.04, respectively). One noteworthy point drawn from Bonferroni/Dunn *posthoc* tests is that unaccented /#ba/ shows an increase in displacement before a higher boundary when /#ba/ is *unaccented* (IP>Wd, p<0.01).

V_1-to-#C_2 (transboundary) lip closing movement. As summarized in Table 5, there is systematic boundary-induced kinematic variation in all measured kinematic variables. V_1-to-C_2 lip closing gestures at a higher prosodic boundary show a progressive increase in displacement, total movement duration,

time-to-peak-velocity and deceleration duration, but a progressive decrease in peak velocity. This pattern holds for both vowels.

Table 4. Summary of boundary effects on C-to-V lip opening movements. The results of posthoc tests (p<0.01) is provided when there is a main effect.

Kinematic measures	domain-final (C_1-to-V_1#)		domain-initial (#C_2-to-V_2)	
	/ba#/	/bi#/	/#ba/	/#bi/
Displacement	$F_{[1.2,6.0]}=0.71^{n.s.}$ (IP=ip)>Wd (when unaccented)	$F_{[1.1,5.8]}=6.40^{*}$ (IP=ip)>Wd	$F_{[1.3,6.9]}=1.71^{n.s.}$ —	$F_{[1.1,5.6]}=0.99^{n.s.}$ —
Total Duration	$F_{[1.1,5.6]}=23.85^{**}$ IP>ip>Wd	$F_{[1.6,8.0]}=66.14^{**}$ IP>ip>Wd	$F_{[1.8,9.3]}=16.53^{**}$ IP>(ip=Wd)	$F_{[2,10]}=9.91^{**}$ IP>(ip=Wd)
C_1ONS-To-V_1PKVEL	$F_{[1.2,5.9]}=8.63^{*}$ (IP=ip)>Wd	$F_{[1.8,9.1]}=49.29^{**}$ IP>ip>Wd	$F_{[1.9,9.4]}=25.85^{**}$ (IP=ip)>Wd	$F_{[2,10]}=17.02^{**}$ IP>ip, IP>Wd
V_1PKVEL-To-V_1TARG	$F_{[2,10]}=8.05^{**}$ IP>ip>Wd	$F_{[1.3,6.6]}=38.79^{**}$ IP>ip>Wd	$F_{[1.1,5.8]}=2.02^{n.s.}$ —	$F_{[1.1,5.9]}=1.01^{n.s.}$ —
Peak Velocity	$F_{[1.1,5.9]}=2.10^{n.s.}$ —	$F_{[1.2,6.0]}=3.02^{n.s.}$ —	$F_{[1.3,6.5]}=1.97^{n.s.}$ —	$F_{[1.2,6.1]}=1.02^{n.s.}$ —

Table 5. Summary of boundary effects on V_1-to-#C_2 lip closing movements. The results of posthoc tests (p<0.01) is provided when there is a main effect.

	/ba#/	/bi#/
Displacement	$F_{[1.3,6.6]}=9.05^{*}$ (IP=ip)>Wd	$F_{[1.4,7.1]}=15.777^{*}$ IP>ip>Wd
Total Duration	$F_{[1.4,7.2]}=20.018^{**}$ IP>ip>Wd	$F_{[1.2,6.6]}=66.654^{**}$ IP>ip>Wd
V_1ONS-To-C_2PKVEL	$F_{[1.4,6.7]}=15.289^{**}$ IP>ip>Wd	$F_{[1.2,6.1]}=55.448^{**}$ IP>ip>Wd
C_2PKVEL-To-C_2TARG	$F_{[2,10]}=35.001^{**}$ IP>ip>Wd	$F_{[2,10]}=37.049^{**}$ IP>ip>Wd
Peak Velocity	$F_{[1.7,8.7]}=32.754^{**}$ IP<ip<Wd	$F_{[2,10]}=5.978^{**}$ IP<ip<Wd

Dynamical aspects of boundary effects. As was the case for Accent effects, the boundary-induced kinematic variations are not fully accounted for by any single dynamical parameter setting. First let's consider lip opening movements. There are some close cases in which the kinematic patterns suggested by ANOVA match the shrinking account, showing the requisite larger and longer movement with no change in peak velocity, especially for domain-final kinematic patterns. The shrinking account for *domain-final* cases appears to be further supported by kinematic relationships: (a) a close relationship between duration and displacement/velocity ratio with a remarkable separation among boundary types ($R^2 = 0.82$ to 0.89 for /ba#/; $R^2 = 0.85$ to 0.97 for /bi#/), which matches the idealized pattern of a change in shrinking (Fig. 1h); (b) a close relationship between total movement duration and time-to-peak-velocity ($R^2 = 0.73$ to 0.78 for /ba#/; $R^2 = 0.86$ to 0.96 for /bi#/). (Note that although Byrd and Saltzman (1998) used this temporal relationship as an index of the degree of stiffness, the close relationship between total movement duration and time-to-peak-velocity also supports the shrinking account because the re-scaling involves a proportional change between the two measures.) However, a close examination of the relationship between peak velocity and displacement, as shown in the left panel of Fig. 3 reveals that the shrinking account is not an absolute fit to the observed pattern, not even domain-finally. If the kinematic pattern were due to a pure change in shrinking, datapoints would be horizontally scattered (see the idealized picture in Fig. 1g), showing a distinct separation among boundary types, which is not what we observe in the figure.

Domain-initially, the evidence for the shrinking account becomes even less clear because of quite a substantial overlap among points belonging to different boundary types (figures not shown). Again, the relationship between peak velocity and displacement (the right panel of Fig. 3) reinforces this by showing substantial overlap among datapoints, rather than the idealized horizontal distribution of datapoints.

Instead, an interesting pattern emerges from the plots in Fig. 3, especially for domain-final cases: the data points are generally scattered diagonally with the datapoints for IP clustering beneath the regression line to the right, and the datapoints for Wd clustering above the regression line to the left. This pattern appears to indicate that some kind of complicated, yet, systematic kinematic mechanism is involved in marking prosodic boundaries, though not accounted for by any single dynamical parameter setting.

Now, let us move on to lip closing (V_1-to-C_2) movements. As seen earlier, V_1-to-C_2 closing gestures at a higher boundary show progressive increase in displacement, total movement duration, time-to-peak-velocity and deceleration duration, but a progressive decrease in peak velocity. This holds for both vowels. The pattern of a larger, longer, and *slower* movement does not single out any particular dynamical parameter as an underlying mechanism. For example, while the patterning of the longer duration with a lowered peak velocity favors the stiffness account, the systematic variation in displacement requires a further dynamical mechanism which cannot be pinpointed here.

Relationships between kinematic measures show evidence that might favor the stiffness account to some extent. The longer and slower movement for V_1-to-C_2 lip closing gestures at a higher prosodic boundary might be accounted for by a decrease in stiffness, as evident in: (a) a close relationship between total movement duration and time-to-peak-velocity (R^2=0.88–0.97); (b) a close relationship between duration and displacement/velocity ratio (R^2=0.88–0.97) with datapoints for a higher prosodic boundary gathering towards the upper right corner of the regression space (bearing resemblance to the idealized picture in Fig. 1b); and (c) datapoints for a higher prosodic boundary being scattered in the lower side of the regression space that relates peak velocity and displacement (again bearing some resemblance to the idealized picture in Fig. 1a, but for the actual plots, see Cho, 2002). While these results indicate apparent temporal aspects that support the stiffness account to some extent, however, the systematic change in displacement (the higher the prosodic boundary, the larger the movement) adds a great deal of dynamic complexity, which again makes it difficult to pinpoint a unified dynamical account.

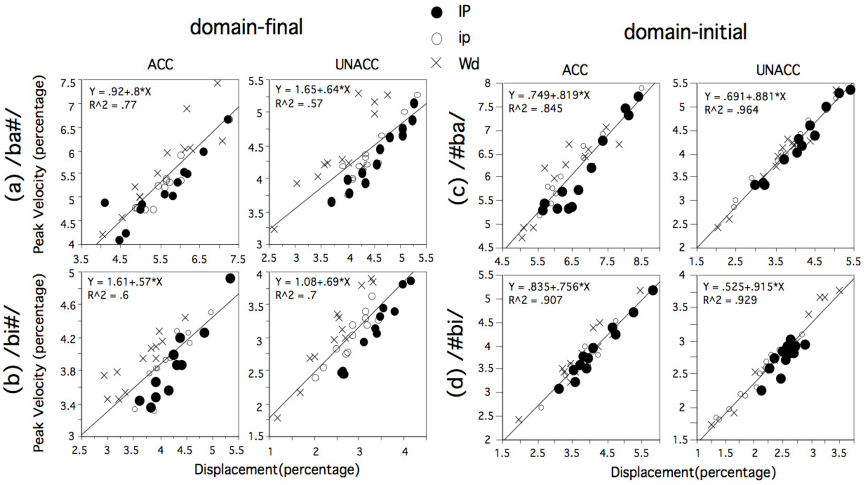

Figure 3. Effect of Boundary Type on relationship between peak velocity and displacement for lip opening gestures, by vowel Type and Accent.

4. General Discussion

Accent-driven kinematic characteristics. In this study, we found that the *lip opening* gesture under accent is associated with an increase in almost all measured kinematic variables including displacement, total movement duration, time-to-peak-velocity, deceleration duration, and peak velocity, regardless of whether it is domain-initial or domain-final. This indicates that the accent-induced articulatory strengthening can be further characterized with a larger, longer, and faster lip opening movement. This result is consistent with findings for jaw opening gestures under accent reported in the literature (as in English *put* reported in de Jong 1991; and in *Pope* and *pipe* in Fowler 1995), but not with those reported in Beckman, et al. (1992) who found that under accent, the jaw opening gesture is associated with an increase in duration and displacement without a substantial increase in peak velocity. The acc/uacc differences are also in line with those coming from lexical stress (Kelso, et al. 1985 for jaw and lower lip movements for reiterant /ba/).

With respect to the *lip closing* (V_1-to-C_2) gesture, one of the significant findings is that the V_1-to-C_2 lip closing gesture is influenced not only by postboundary (V_2) accent but also by preboundary (V_1) accent. Some measured variables are affected primarily by preboundary accent and some, by postboundary accent, while yet others are affected by both. For example, spatial displacement is significantly affected by both V_1 Accent and V_2 Accent, such that lip closing displacement is larger for ACC than for UNACC, regardless of whether accent comes from the preboundary or the postboundary syllables. As for durational variation, the preboundary accent affects primarily the first durational component (time-to-peak-velocity) of the movement duration, whereas the postboundary accent affects only the second component (deceleration duration) of the movement duration. Finally, while peak velocity is consistently influenced by preboundary (V_1) accent such that it is higher for ACC than for UNACC, there is no consistent effect of postboundary (V_2) accent on peak velocity. In short, although there are some compounding effects of Accent arising from both sides of the boundary, the effects of both V_1 and V_2 Accents converge on a larger and longer lip closing movement, while a faster movement comes primarily from V_1 Accent. This is generally consistent with accent-induced kinematic characteristics for the lip opening gesture.

As an aside, a noteworthy point concerns whether the V_1-to-C_2 lip closing gesture should be considered solely as a postboundary phenomenon. Byrd and Saltzman (1998) define the V_1-to-C_2 lip closing gesture as being postboundary because it is activated in order to form a lip constriction for $\#C_2$

which belongs to the postboundary syllable. According to this account, the V_1-to-C_2 lip closing gesture should perhaps be influenced only by postboundary (V_2) accent. However, the fact that some measured kinematic variables are affected only by preboundary accent and some only by postboundary accent suggests that kinematic variation for the V_1-to-C_2 lip closing gesture may be better defined as a *transboundary* phenomenon rather than as a postboundary phenomenon. From the results, we can make a generalization that the V_1-to-C_2 lip closing gesture can be thought of as consisting of two components, with the timepoint of peak velocity as a landmark. First, the articulation during the time course from V_1 onset to the peak velocity landmark may be characterized as a preboundary phenomenon which is governed by the preboundary accent. Second, the articulation during the time course from the peak velocity landmark to the C_2 target attainment may be thought of as a postboundary phenomenon which is governed by the postboundary accent. (Note also that if we apply this transboundary nature of the C_2 lip closing gesture to the framework of syllable structure, we can further posit that the C_2 lip closing gesture is "ambisyllabic" in that the first half of it belongs to the preceding syllable and the second half of it to the following syllable.)

Boundary-driven kinematic characteristics. The one obvious kinematic characteristic for both lip opening and closing gestures at edges of prosodic domains is that they are consistently longer, but this time, not necessarily faster. However, there is an inconsistent boundary effect on spatial displacement in lip opening and closing gestures. The larger displacement was found consistently for the cross-boundary V_1-to-C_2 lip closing gesture, showing a progressive increase in displacement as the prosodic boundary moves up in the prosodic hierarchy. This is consistent with results reported in Byrd and Saltzman (1998). On the other hand, domain-edge lip opening gestures show some interaction between Accent and Vowel Type. (As pointed out by a reviewer, the kinematic difference between lip opening and closing gestures may be in part due to some physiological reasons: there is lip compression at the end of a closing gesture which is likely to change articulatory patterns.)

As for the *domain-final* lip opening gestures, /bi#/ shows an increase in displacement before a higher prosodic boundary, but /ba#/ shows such an effect only when it is *unaccented*. Similarly, for the *domain-initial* lip opening gesture, there is an increased displacement after a higher boundary only for *unaccented* /ba#/. There is thus an increased displacement at a higher boundary at least when the target gestures are unaccented. This is presumably because of some sort of ceiling effect due to accent, that is, when gestures are accented, articulation is already expanded such that an expanded articulation would not leave much room for an additional articulatory expansion from boundary type.

At first glance, this result appears to be consistent with previous findings (e.g., Edwards, et al., 1991) whereby an expanded jaw opening displacement was found domain-finally, only when the gestures being compared are *unaccented*. Recall, however, that some of our results show an expanded lip opening displacement when *accented* (e.g., final /bi#/ and transboundary /a#b/ and /i#b/), suggesting that boundary-induced spatial expansion is not limited to the unaccented gestures only. (However, it should be noted that the difference between the present study and previous studies may be due to the articulators that have been examined. See below for discussion on limitations of lip kinematics.)

All in all, the results suggest that there is some sort of articulatory strengthening as evident by the longer and sometimes larger lip opening and closing gestures at a higher prosodic boundary. However, the boundary-induced strengthening pattern is somewhat different from that arising from accent in that the latter is associated with a faster movement whereas the former is not. Further, this pattern, especially the longer opening movement duration, is found not only in domain-final but also in domain-initial positions. As discussed above, while Byrd and Saltzman (1998) did not consider domain-initial lip opening gestures (thus, for example, it may not undergo lengthening), the present study suggests that the domain-initial lip opening gesture has temporal characteristics much the same as the domain-final lip opening gesture.

4.1. Can prosodically-driven kinematic variations be modeled by a particular dynamical parameter setting?

Accent-driven kinematic variations. Some researchers (e.g., Edwards, et al. 1991; Harrington, et al. 1995) have proposed that an intergestural timing mechanism underlies accent-induced kinematic variation in jaw opening and closing gestures. However, when the present kinematic findings regarding lip opening gestures were considered, no single dynamical parameter setting could be singled out as the underlying mechanism. For the lip opening gesture, the longer and larger movement pattern suggested that, if anything, a change in both stiffness and target was the more probable account for accent-induced differences, with a change in intergestural timing as the least likely mechanism. (Note that findings reported by de Jong (1991) and Fowler

(1995) also suggested that if anything an increase in target is the most likely source for an increased displacement.) For the cross-boundary V_1-to-C_2 lip closing gesture, the effect of preboundary (V_1) accent shows a pattern similar to the effect of accent on the opening gesture, favoring no dynamical account. The longer and larger articulation (with no change in peak velocity and time-to-peak-velocity) due to postboundary accent for /i#b/ seems to be ascribable to a change in intergestural timing. Again, however, relationships between various kinematic variables did not support this, weakening the intergestural timing account.

What emerges from the data is then that no single dynamical mechanism can account for accent-induced kinematic variations, contrary to what has previously been assumed among researchers who have attempted to characterize prosodically-conditioned kinematic variations in terms of a mass-spring dynamical parameter setting.

Boundary-driven kinematic variations. As was the case for Accent effect, the boundary-induced kinematic variations were not fully accounted for by any single dynamical parameter setting. If we consider only temporal kinematic measures, namely, the total movement duration and time-to-peak-velocity, as Byrd & Saltzman (1998) did, the boundary-induced durational difference is likely ascribable to a change in stiffness, given the proportional change in the total movement duration and time-to-peak-velocity as a function of prosodic boundary. However, when we consider additional kinematic measures, the stiffness hypothesis is seriously undermined. For instance, when peak velocity (which was not included in Byrd & Saltzman) is considered, only the lip *closing* gesture shows a slower movement (with lowered peak velocity) at a higher prosodic boundary, favoring the stiffness account, whereas no change in peak velocity in the case of the lip *opening* gestures weakens the stiffness account. Moreover, when the variation in displacement is figured in, it becomes even more obvious that a change in stiffness is not the only dynamical mechanism underlying the boundary-induced longer, larger, and sometimes slower movement.

Here, it is worthwhile noting two possible sources of variation in displacement. Both Byrd & Saltzman and the present study have measured the displacement of the lip closing gesture by differentiating the lip opening maxima and minima. However, as pointed out by Goldstein (p.c.), the displacement in lip opening in $V_1\#C_2$ may vary not only due to a change in the target of the lip closing gesture but also due to a change in the value of the Lip Aperture at the onset of the gesture associated with the preceding vocalic gesture. This becomes clearer with the results of the present study regarding accent-induced variation in displacement. It was found that the V_1-to-C_2 lip closing gesture was associated with an increase in displacement when either V_1 or V_2 was accented. It is therefore possible that when V_1 was accented, the increased displacement was mainly due to the more extreme opening value at the onset of lip closing while the increased displacement due to accented V_2 was primarily attributable to the more extreme target value obtained during C_2 constriction. In other words, as Cho (2002) discussed, the reason for effects of both V_1 Accent and V_2 Accent is presumably because the maximum Lip Aperture for V_1 is significantly larger for $V_{1\text{ACC}}$ than $V_{1\text{UNACC}}$ ($p<0.01$), and the minimum Lip Aperture for C_2 is significantly smaller for $V_{2\text{ACC}}$ than for $V_{2\text{UNACC}}$ ($p<0.01$).

One might then question whether measuring the spatial difference (displacement) between the onset and the offset of the V_1-to-C_2 movement adequately reflects the target (gestural amplitude) of the relevant dynamical system (here, the lip closing gesture). For example, the increased displacement at a higher prosodic boundary found in this study may not exclusively reflect the change in target in the dynamical system. Cho (2002) indeed reported that Lip Aperture maxima for preboundary V_1 were generally larger at a higher prosodic boundary than at a smaller prosodic boundary. Thus, it requires caution to interpret boundary-induced kinematic variation in displacement in terms of a dynamical parameter setting, for both the present study and Byrd & Saltzman 1998.

Some discussion on the relationship between kinematics and dynamics . With all these in mind, let us return to the issue of how the kinematic results can be accounted for in the framework of dynamics. One might raise a rather fundamental question about the validity of the current mass-spring dynamical model. If the dynamical model were assumed to predict that modification to a single dynamical parameter is the only way to control kinematics, then the failure to single out any particular dynamical parameter setting would suggest that the current mass-spring dynamical model is not adequate to account for the prosodically-induced kinematic patterns. With respect to accent-induced kinematic pattern, Fowler (1995) indeed proposed that gestural behaviors under accent may not be best described in terms of dynamical parameter settings, but rather they are most consistent with the “global effect” hypothesis that stress consists of a global increase in *production effort* in order to maximize prominence in the stressed syllable. Such prominence maximization can then be obtainable simply by the larger, longer, and faster lip opening and closing movements, as found in this study.

Alternatively, speech mechanisms may not be as simple as has been assumed by researchers who adopt the mass-spring dynamical model in explaining certain speech phenomena. The observed data could be explained under a mass-spring dynamical model, if we further explore the possibility that more than one dynamical parameter governs the accent-induced kinematic patterning. For example, from the present study, one might infer that both stiffness and target changes govern the lip opening movements under accent, and the lip closing movement under V_1 accent, whereas changes in both stiffness and intergestural timing likely underlie the lip closing movement under V_2 accent. Likewise, the consistently larger displacement in the lip closing movement at a higher prosodic boundary can be dealt with by either the target or the intergestural timing parameters in combination with the stiffness parameter that accounts for the observed temporal aspects. Further, we cannot entirely reject the possibility that all the dynamical parameters are interactively influential on kinematic realizations with different degrees of effect, such that breaking down such compounding effects into individual dynamical parameter settings would be extremely difficult without fine-grained computational modeling on ample empirical data.

Finally, there is another caveat interpreting the kinematic data of the present study in dynamical terms. In this study, following Byrd & Saltzman (1998), the lip opening and closing movements (Lip Aperture) have been assumed to be regulated by a single dynamical regime (gesture). While it is reasonable to assume that lip closing is controlled by a single dynamical system (i.e., a consonantal lip closing gesture), it is less clear whether the lip opening (e.g.,. C_1-to-V_1 and C_2-to-V_2 movements) is indeed modulated by a single dynamical gesture (Goldstein, p.c.). Lip opening movements are usually associated with a vocalic gesture which may regulate tongue task variables primarily, and Lip Aperture may be influenced by not only the tongue movement but also the action of the jaw which accompanies it. Therefore, the failure to interpret lip opening kinematics in terms of dynamics may be attributable in part to such articulatory complexity associated with the lip opening during the vocalic movement.

At the very least, however, the findings in the present study motivate future studies to look for not only the complexity of dynamical parameter settings but also articulatory complexity associated with a single gesture, rather than seeking what particular dynamical parameter setting ‘best’ matches speech phenomena. However, even if practicing linguists can develop such a complicated model (building on the currently available dynamical model) which can adequately describe all the complex kinematic patterns as presented in the present paper, it will still be interesting to see how such a complex dynamical system is learned in the course of the language acquisition.

The π-gesture. Another way of characterizing boundary-adjacent kinematic variation is suggested by Byrd and Saltzman (Saltzman 1995, Byrd, et al. 2000; Byrd 2000; Saltzman & Byrd 2000; Byrd & Saltzman 2003): there might be abstract, non-tract variable prosodic boundary gestures that are governed by prosodic constituency in a mass-spring dynamical model. The so-called 'π-gesture' was hypothesized initially to affect stiffness in tract variable articulatory gestures over its activation period, roughly in proportion to the strengths of the boundary: the larger the prosodic boundary, the less stiff the articulatory gestures in the vicinity of the boundary. In Byrd & Saltzman (2003), the stiffness modulation approach was replaced with the clock-slowing modulation approach: the π-gesture locally slows the clock that controls the timecourse of gestural activation. In this framework, boundary-induced temporal variation can be interpreted as a change in clock-slowing under the influence of the π-gesture. The temporal activation interval of the π-gesture overlaps with the activation interval of articulatory gestures adjacent to prosodic boundaries, such that the boundary-adjacent articulation lengthens in proportion to degree of the π-gesture's strength, which is again roughly proportional to level of prosodic boundary.

Degree of lengthening is also influenced by the temporal extent of the π-gesture. In an earlier model of π-gesture, Byrd (2000) suggested that the π-gesture's domain of influence is local at edges of prosodic domains – i.e., "only the constriction gestures within the π-gesture's temporal field of activation are directly affected, not gestures remote from the phrasal boundary (p. 14)." Thus, Byrd hypothesized that for the sequence $C_1V_1\#C_2V_2$, articulations that are closest to the prosodic boundary are most influenced by the π-gesture, resulting in the maximal elongation. In the present study, however, it was found that not only V_1-to-C_2 movement (which is the closest to the prosodic juncture) but also C_1-to-V_1 and C_2-to-V_2 movements were all significantly affected by boundary type. Of course, it is likely that the lengthening of C_1-to-$V_1\#$ comes from the effect of the π-gesture on $V_1\#$ and the lengthening of $\#C_2$-to-V_2, from the effect of the π-gesture on C_2. It is also possible that articulations for the rather remote C_1 and V_2 are still within the activation field of the π-gesture but presumably with somewhat reduced degree of the π-gesture's influence, under the assumption that the π-gesture's strength tapers out towards edges of its temporal activation interval. However, it is not entirely clear what are the exact mechanisms that underlie lengthening of gestures that are not immediately next to a prosodic boundary. Byrd and

Saltzman (2003) explain that "its [π-gesture's] effect will be felt on any of the gestures with which it is coarticulated; under the assumption that the π-gesture is anchored to the prosodic juncture, these will be gestures closest to the phrase edges." Specifically, it is hoped that future studies provide us with more information not only about the precise temporal extent of the π-gesture, but also about its relationship with the declining nature of the π-gesture's strength towards the edges of the activation interval.

Another issue regarding the π-gesture model is whether the π-gesture influences degree of spatial magnitude directly or not. The available information with respect to variation in spatial magnitude comes from simulations (Saltzman and Byrd 2000, Byrd and Saltzman 2003) which demonstrate that a clock-slowing implementation of π-gestures may entail variation in displacement associated with domain-initial consonant constriction. For example, for the domain-initial consonant-vowel constriction sequence, the π-gesture initiates the CV constriction sequence (rather than intervening it) such that under the influence of π-gesture, the consonantal constriction will get not only longer but also overlap less with the following vowel gesture. In the current model, the decreased overlap (or less truncation) between the consonantal gesture and the following vocalic gesture accounts for the domain-initial strengthening phenomenon – i.e., the increase in gestural amplitude associated with domain-initial consonants (e.g., Fougeron and Keating 1997; Cho and Keating 2003). It remains to be seen whether this model would be able to account for the full range of results presented here and elsewhere, and again whether it could provide a simple and unified theory about the prosodically-driven systematicity in speech production.

4.2. Enhanced consonant-vowel contrasts at domain-edges

One of the central issues with respect to boundary-induced kinematic variation is whether it is a linguistically significant phenomenon. It has been suggested in the literature (Fougeron and Keating 1997; Hsu and Jun 1998; Fougeron 2001) that expanded #CV or V#C displacement adjacent to a prosodic boundary would serve as an articulatory signature for marking that prosodic boundary. The present results are generally supportive of this proposal. In particular, the V#C lip closing gesture shows the most robust boundary effect on displacement with a pattern of IP>ip>Wd. A similar result was found for domain-final (CV#) lip opening gesture, whereas the domain-initial (#CV) lip opening gesture did not show a consistent effect. This is compatible with Fougeron and Keating's observation that domain-initial consonantal strengthening, as measured by linguopalatal contact, induces a greater V#C displacement at edges of prosodic domains, while such an effect is less evident in degree of #CV displacement. This observation is reinforced by kinematic data reported in this study. Further, the results presented in this study show that even the domain-final CV# displacement is expanded at higher prosodic boundaries, which Fougeron and Keating did not find in their EPG data. Overall, we can infer that contrasts between consonants and vowels are enhanced at edges of prosodic domains (syntagmatic contrast enhancement) via an increase in displacement adjacent to a prosodic boundary, which can be seen as the articulatory manifestation of prosodic structure. (See Cho and Jun (2000), Cho and McQueen (2005) for discussion on domain-initial strengthening in terms of enhancement of distinctive features.) Recent research has begun to investigate the role of domain-initial consonantal strengthening in lexical segmentation in English (McQueen and Cho 2003; Cho, McQueen and Cox, in press), showing that the acoustic consequences of initial strengthening facilitate word recognition.

4.3. Conclusion

The present study has investigated how segmental phonetic realizations are conditioned by various prosodic factors by examining kinematic variations in accented syllables, domain-initial, and domain-final syllables. While previous studies have looked at these locations separately, the present study differs from them in that it examined all these locations concurrently. Crucially, each of the three prosodically important locations showed distinctive kinematic patterns that can distinguish itself from others. Several major points have emerged. First, accent-induced articulatory strengthening can be characterized by larger, longer, and faster lip opening and closing movements. That is, when accented, vowel movements are simply bigger in all ways – in distance, time, and speed. Second, unlike accent-induced strengthening, boundary-induced strengthening effects are evident in longer, but this time not necessarily faster, articulation in both domain-initial/final positions. The spatial expansion is found quite consistently at the domain edge when the gestures are *unaccented,* and more consistently for cross-boundary lip closing gestures regardless of accent. Further, temporal characteristics are similar for the domain-final and initial lip opening gestures (i.e., longer duration with no change in velocity), though the domain-final gesture is longer than domain-initial one. In short, all of the three prosodically important locations show strengthening effects with generally longer and larger movements in common, but they differ primarily in velocity: faster for accented gestures, no change for domain-initial/final lip opening gestures, and slower for cross-boundary lip closing gestures.

Finally, the results regarding movement kinematics suggest that speech mechanisms are more complex than has generally been assumed. These results challenge the theories previously advanced in the framework of a mass-spring gestural model. It was proposed that in order to account for prosodically-conditioned kinematic patterns in the framework of a mass-spring gestural model, at the very least one should look for a combination of settings for multiple dynamical parameters, rather than seeking one particular dynamical mechanism governing kinematic patterns arising from each prosodic condition. Alternatively, the best solution to the problem might be to find a simple and unified dynamical theory (not necessarily in the framework of a mass-spring gestural model) which can model the prosodically-driven systematicity in a way that is both descriptively and explanatorily adequate.

This study suggests that phonetic realization is governed by high level prosodic conditions, and that prosodically-conditioned kinematic patterns, in turn, manifest high level prosodic structures. Furthermore, the systematic phonetic variation conditioned by prosodic structure should be taken more seriously into account in developing linguistic theories, especially in modeling speech production and speech perception. It is ultimately hoped that this study will contribute to theories of the phonetics-prosody interface, making progress towards gaining better insight into prosodically-driven speech phenomena.

Acknowledgments

I thank Pat Keating, Sun-Ah Jun, Peter Ladefoged, Bruce Hayes, Jody Kreiman, Adam Albright, Elizabeth Johnson and James McQueen for their comments and suggestions. Special thanks goes to Louis Goldstein and two other anonymous reviewers for critical and constructive comments. This work was supported in part by the NSF doctoral improvement grant (BCS-0001716) to the author and Pat Keating.

References

Beckman, M.E.
1996 The parsing of prosody. *Language and Cognitive Processes* 11: 17–67.

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

Beckman, M.E. and G.A. Elam
1997 Guidelines for ToBI labelling, version 3.0. Unpublished Manuscript. Ohio State University.

Beckman, M.E. and J. Pierrehumbert
1986 Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Browman, C. and 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 the Physics of Speech*, 341–376. Cambridge: Cambridge University Press.

Byrd, D.
2000 Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57: 3–16.

Byrd, D. and E. Saltzman
1998 Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26: 173–200.
2003 The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31: 149–180.

Byrd, D., A. Kaun, S. Narayanan and E. Saltzman
2000 Phrasal signatures in articulation. In: M. Broe and J. Pierrehumbert (eds.), *Acquisition and the Lexicon: Papers in Laboratory Phonology V*, 70–88. Cambridge, UK: Cambridge University Press.

Cho, T.
2002 *The Effects of Prosody on Articulation in English*. New York and London: Routledge.
2004 Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32: 141–176.
2005 Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America* 117(6): 3867–3878.

Cho, T. and J. McQueen
2005 Prosodic influences on consonant production in English: Effects of prosodic boundaries, phrasal accent and lexical stress. To appear in *Journal of Phonetics* 33(2): 121–157.
Cho, T., J. McQueen and E. Cox
in press Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics.*
Cho, T. and P. Keating
2001 Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155–190.
Cho, T. and S.-A. Jun
2000 Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society 36* (CLS 36): 31–44.
De Jong, K.
1991 The Oral Articulation of English Stress Accent. Ph.D. Dissertation, Ohio State University.
1995 The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491–504.
Edwards, J.E., M.E. Beckman and J. Fletcher
1991 The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89: 369–382.
Fougeron, C.
2001 Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29: 109–135.
Fougeron, C. and P.A. Keating
1997 Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101: 3728–3740.
Fowler, C.A.
1995 Acoustic and kinematic correlates of contrastive stress accent in spoken English. In: F. Bell-Berti and J.J. Raphael (eds.), *Producing Speech: Contemporary Issues: For Katherine Safford Harris*, 355–373. Woodbury: American Institute of Physics.
Harrington, J., J. Fletcher and C. Roberts
1995 Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics* 23: 305–322.
Hawkins, S.
1992 An introduction to task dynamics. In G. Docherty and D.R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* 9–25. Cambridge: Cambridge University Press.

Hsu, C.-S. and S.-A. Jun
1998 Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics* 96: 69–89.

Jun, S.-A.
1993 The Phonetics and Phonology of Korean Prosody. Ph.D. Dissertation, Ohio State University.
1998 The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15: 189–226.

Keating, P.A., T. Cho, C. Fougeron and C. Hsu
2003 Domain-initial strengthening in four languages. In: J. Local, R. Ogden, and R. Temple (eds.), *Papers in Laboratory Phonology IV: Phonetic Interpretation,* 145–163. Cambridge, UK: Cambridge University Press.

Kelso, J.A.S., E. Vatikiotis-Bateson, E. Saltzman and B. Kay
1985 A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling. *Journal of the Acoustical Society of America* 77: 266–280.

McQueen, J. and T. Cho
2003 The use of domain-initial strengthening in segmentation of continuous English speech. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*. Barcelon, Spain 2003.

Pierrehumbert, J. and M.E. Beckman
1988 *Japanese tone structure*. Cambridge, Mass.: MIT Press.

Pierrehumbert, J. and D. Talkin
1992 Lenition of /h/ and glottal stop. G. Docherty and D.R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* 90–117. Cambridge: Cambridge University Press.

Saltzman, E.
1995 Intergestural timing in speech production: Data and modeling. In *Proceedings of the XIII International Congress of Phonetic Sciences 2*: 84–91.

Saltzman, E. and D. Byrd
2000 Demonstrating effects of parameter dynamics on gestural timing. Presented at the May 2000 meeting of the Acoustical Society of America, Atlanta, Georgia. *Journal of the Acoustical Society of America* 107(5,2): 2904.

Shattuck-Hufnagel, S. and A.E. Turk
1996 A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25: 193–247.