Journal of Phonetics 44 (2014) 96–109

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures

Taehong Cho [a,*], Yeomin Yoon [a,b], Sahyang Kim [c]

[a] Hanyang Phonetics and Psycholinguistics Lab, Department of English Language and Literature, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Republic of Korea
[b] Audio Research Center, TranSono Inc., Seoul, Republic of Korea
[c] Department of English Education, Hongik University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 11 March 2013
Received in revised form
4 February 2014
Accepted 20 February 2014
Available online 18 March 2014

Keywords:
CV coordination
Intergestural timing
Temporal stability
Prosodic boundary
Syllable structure
Domain-initial strengthening
Prosodic strengthening
Pi-gesture
Gestural coupling
EMA
Korean

ABSTRACT

An articulatory study was conducted to explore effects of prosodic boundary and syllable structure on temporal realizations of /ma/ in C♯V vs. ♯CV in Korean (where '♯' denotes an Intonational Phrase or a Word boundary). The vocalic gesture underwent boundary-induced lengthening more in C♯V̲ than in ♯CV̲, implying that the boundary effect is largely localized to the initial element whether consonantal or vocalic. CV coordination patterns were temporally neutralized between ♯CV and C♯V in the phrase-internal Word boundary condition, showing a possible 'resyllabification' of 'C' with the following vowel in C♯V in the articulatory temporal measures taken in the present study. It was suggested that CV gestures in C♯V, whose phasing relationship has to be determined postlexically, reorganize temporally in an in-phase coupling mode just like the way CV gestures are phased in ♯CV. Finally, while there was leftward shifting of the consonantal gesture in C♯V with some temporal variability across an IP vs. a Word boundary, intergestural timing in ♯CV remained invariant regardless of boundary strength. But the most stable temporal pattern was observed with an IP boundary in ♯CV, interpretable as an important temporal characteristic of domain-initial strengthening. Some of these results were further discussed in terms of their implications for the theory of π-gesture and the gestural coupling model of syllable structure.

1. Introduction

One of the important theoretical assumptions in Articulatory Phonology (Browman & Goldstein, 1990, 1992, 2000; Goldstein & Fowler, 2003) is that linguistic information such as phonological contrast and lexical distinction may be encoded by differential intergestural timings—i.e., how gestures are temporally coordinated with each other. For example, a subtle change in timing between the consonantal release gesture for a stop and the laryngeal gesture that initiates vocal fold vibration may result in a category change from voiced to voiceless (or vice versa) especially when the change occurs around the category boundary (see Cho & Ladefoged, 1999, for VOT defined in terms of intergestural timing). Some degree of stability in intergestural timing must therefore be guaranteed to maintain the underlying phonological and lexical information in a given language. But at the same time, some degree of variability must also be allowed, given that speech timing by nature is variable in connected speech, and that it is modified by various other factors (e.g., intrinsic factors such as the tongue height for vowels and the place of articulation for consonants, and extrinsic factors such as segmental and suprasegmental contexts; see Cho (to appear) for a review on speech timing). Understanding the sound structure of language as a dynamical system therefore requires understanding not only how individual gestures are temporally realized, but also how their coordination and temporal stability (or variability) vary as a function of multiple linguistic and extralinguistic factors.

In the present study, we continue to explore the dynamical nature of speech timing by investigating how two interrelated factors, prosodic boundary and syllable structure, interactively influence the temporal realization of CV gestures in Korean. The goal of this study is two-fold: first, to understand how the temporal realization of individual consonantal and vocalic gestures is influenced by prosodic boundary strength (an Intonational Phrase (=IP) boundary vs. a Prosodic Word (=Wd) boundary) in different syllable structure conditions (C♯V vs. ♯CV); and second, to understand how timing between the two gestures (i.e., CV coordination) and their temporal stability are modified by an interaction between prosodic boundary and syllable structure. In the remaining part of this section, specific research questions will be introduced along with some theoretical motivations.

* Corresponding author. Tel.: +82 2 2220 0746; fax: +82 2 2220 0741.
E-mail addresses: tcho@hanyang.ac.kr, taehong.cho@gmail.com (T. Cho), yoon.yeomin@gmail.com (Y. Yoon), sahyang@hongik.ac.kr (S. Kim).

0095-4470/$ - see front matter
http://dx.doi.org/10.1016/j.wocn.2014.02.007

It has been well documented in the literature that temporal realizations of consonantal and vocalic gestures are influenced by prosodic boundary, such that they are produced with longer articulatory durations at a larger prosodic boundary (e.g., an IP boundary) than at a smaller prosodic boundary (e.g., a Wd boundary) (e.g., Byrd, Krivokapić, & Lee, 2006; Byrd & Saltzman, 2003; Cho, 2005, 2006, 2008; Cho & Keating, 2001, 2009; Cho & McQueen, 2005; Edwards, Beckman, & Fletcher, 1991; Fougeron, 2001; Krivokapić & Byrd, 2012; Tabain, 2003; see Fletcher, 2010 or Cho, 2011, for a review). One of the important questions regarding the boundary-related lengthening effect concerns its scope on domain-initial (postboundary) segments. It has been observed that the postboundary lengthening effect (also known as the domain-initial strengthening effect) is by and large limited to the initial consonant and its effect on the following vowel is null or, if it exits, attenuated (e.g., Barnes, 2002; Byrd et al., 2006; Cho, 2006; Cho & Keating, 2001, 2009; see also Byrd & Choi, 2010 and Bombien, Mooshammer, Hoole, & Kühnert, 2010 for relevant boundary effects on consonant clusters in English and German, respectively). These studies, however, have some limitations. They often considered ♯CV contexts only (but see Fougeron, 2001, for boundary effects on ♯V in French), such that the vowel under investigation was not placed in the strictly initial position. In addition, some of the previous studies (Barnes, 2002; Cho & Keating, 2001, 2009; Cho, Lee, & Kim, 2011) were based only on acoustic data, so we do not have much information available on how vowels are articulatorily manifested in the not-strictly-initial (♯CV) vs. the strictly domain-initial (i.e., ♯V) context.

That said, a question arises as to whether the attenuated or the null lengthening effect on the not-strictly-initial vowel that has often been observed in the acoustic dimension will also be observable in the articulatory dimension. Two alternative possibilities can be considered. On the one hand, it is possible that the null/weak lengthening effect on the acoustic vowel duration is indeed a reflex of the temporal realization of the underlying vocalic gesture which does not undergo domain-initial lengthening in the ♯CV context (i.e., when the vowel is not strictly initial). If this is the case, we should be able to observe differences in vocalic gestures between ♯V and ♯CV contexts. On the other hand, it is also possible that the vocalic gesture may indeed undergo domain-initial strengthening in the ♯CV context, but it is not reflected in the acoustic dimension. Given that *C* and *V* gestures are likely independent from each other, operating on separate articulatory tiers (Browman & Goldstein, 1992; Fowler, 1983), and given that consonantal articulation is lengthened domain-initially, the temporal expansion of consonantal gesture may have an effect of hiding some portion of the vocalic gesture (i.e., with extended CV overlap). As a result, some portion of the vowel in ♯CV may be masked in the acoustic dimension, often failing to show a significant boundary effect on the acoustic vowel duration. If this were the case, the vocalic gestures in both ♯V and ♯CV would show boundary-related lengthening in the articulatory dimension. We test these possibilities by comparing the boundary effects in two syllable structure conditions—i.e., in ♯CV and C♯V. The comparison of the two conditions will also allow us to test whether or not the domain-initial lengthening effect indeed varies between ♯CV and C♯V in such a way that the magnitude of the boundary effect is larger on the strictly initial vowel (in C♯V) in line with the assumption that the domain-initial strengthening may be local to the initial segment whether consonantal or vocalic (Fougeron, 2001; see Cho & Keating, 2009 for a related discussion).

Exploring these possibilities will have further implications for the scope of boundary-induced lengthening in connection with the theory of π-gesture (Byrd, 2000, 2006; Byrd, Kaun, Narayana & Saltzman, 2000; Byrd & Saltzman, 2003; Byrd et al., 2006; Saltzman, 1995).[1] The π-gesture is assumed to be anchored to a prosodic boundary, governing temporal realization of gestures at the juncture, and has its domain of influence which waxes and wanes, so that its effect is strong at the juncture and becomes attenuated in a gradual fashion in both directions (into both preboundary and postboundary segments). A relevant question for the purpose of the present study is then whether the π-gesture influences the temporal realization of the domain-initial vocalic gesture to the same extent regardless of whether the vowel is strictly initial or not (i.e., C♯V vs. ♯CV). Considering the overlapping nature of CV gestures (e.g., Browman & Goldstein, 2000; Fowler, 1983; Löfqvist & Gracco, 1999; Öhman, 1966; Perkell, 1969), and the independent realization of the vocalic gesture in a separate functional articulatory tier in the framework of Articulatory Phonology (e.g., Browman & Goldstein, 1990, 1992), it is reasonable to assume that the vocalic gesture is domain-initial (i.e., immediately adjacent to the prosodic boundary) in both C♯V and ♯CV, although V in ♯CV is not domain-initial in the conventional left-to-right phonetic notation system. In this theoretical framework, the influence of π-gesture on the vowel should in principle remain unchanged, resulting in comparable lengthening effects in ♯CV and C♯V. Testing this possibility will therefore shed light in some detail on how a π-gesture is actually anchored to the articulatory gesture in different syllable structure conditions.

In addition to examining the boundary effect on the vocalic gesture in ♯CV and C♯V conditions, comparison of temporal realizations in ♯CV vs. C♯V will allow us to explore some other interesting questions. It will be examined whether there is any asymmetric boundary effect on the consonantal gesture between the two conditions—i.e., in coda (C♯V) and onset (♯CV) position. As the consonant is immediately adjacent to the boundary in both coda and onset positions, one may predict that both will show comparable lengthening effects. However, considering that the onset/coda asymmetry has often been reported in the literature, showing a longer consonantal duration in coda than in onset position at the word level (e.g., Byrd, 1996; Keating, Wright, & Zhang, 1999), it is possible that the coda-onset asymmetry emerges at a phrase level as well, reflected in the magnitude of boundary-related lengthening—i.e., with a more robust boundary effect on the coda consonant before a boundary than on the onset consonant after it. It will also be interesting to examine whether temporal realization of consonantal and vocalic gestures differs as a function of syllable structure even in the potentially neutralizing (phrase-internal) Word boundary condition. In an acoustic study on French, Fougeron (2007) showed that, although a word-final consonant is 'resyllabified' as an onset of the following syllable across a Word boundary when words are concatenated in running speech (a phenomenon known as *enchaînement*), the resyllabificaiton is not complete, cueing the underlyingly different lexical boundaries. The present study will explore this issue in Korean along the articulatory dimension. (See below for further elaboration of the resyllabification issue in terms of intergestural timing.)

The second general goal of the present study is to investigate how the coordination of consonantal and vocalic gestures and their stability are conditioned by the interaction between syllable structure and prosodic boundary. In Articulatory Phonology, it is hypothesized that intergestural timing is lexically specified, and therefore gestures that belong to the same lexical item are expected to be more stably timed than gestures that do not. While this is an important theoretical assumption of Articulatory Phonology, only a few studies have tested it systematically (e.g., Byrd, 1996; Cho, 2001), and therefore more empirical data are required to understand what it means from the kinematic point of view for two adjacent gestures to belong to the same vs. different lexical items. We therefore continue to explore this question, but this time by including the prosodic boundary factor, so that we can observe how CV coordination and its temporal stability is further modulated by prosodic boundary strength.

The lexical effect, which is dubbed with the syllable structure effect in this study (i.e., ♯CV gestures belong to the same lexical item and C♯V gestures to different lexical items), will be considered in both IP and Wd boundary conditions. But the locus of attention will be on the Wd boundary

[1] In an effort to incorporate boundary-induced lengthening effects into the framework of a gestural model, Byrd and her colleagues proposed that temporal variation in the vicinity of a prosodic boundary can be understood in terms of the influence of a so-called 'π-gesture' (prosody gesture) that is governed by prosodic constituency in the task dynamics model (e.g., Saltzman & Munhall, 1989; see Hawkins, 1992, for an overview for non-specialists). A π-gesture is assumed to modulate the rate of the clock that controls articulatory activation of constriction gestures in dynamical systems, determining their articulatory movement speed.

condition in which the syllable structure difference (♯CV and C♯V) may be potentially neutralized. More specifically, particular attention will be paid to whether CV coordination and its temporal stability (as reflected in the token-to-token variation) manifest any difference at all between C♯V and ♯CV in line with the hypothesis that CV gestures are more stably timed within a lexical item (♯CV) than across a boundary between two lexical items (C♯V). Testing this possibility has further implications for a recently developed gestural coupling model of syllable structure (Goldstein, Byrd, & Saltzman, 2006; Marin & Pouplier, 2010; Mooshammer, Goldstein, Nam, & McClure, 2012; Nam, 2007; Nam, Goldstein, & Saltzman, 2009; Pouplier, 2012). It was proposed that intergestural timing between consonantal and vocalic gestures within a syllable (♯CV and VC♯) can be understood in terms of the intrinsic stable modes (in-phase and anti-phase modes) of a system of coupled oscillators that control timing between gestures in speech production. For example, intergestural timing between C and V gestures is taken to be controlled by two oscillators (roughly speaking, one for C and one for V) that are coupled in an in-phase mode for CV (onset-vowel) coordination (in which two gestures are phased synchronously) and in an anti-phase mode for VC (vowel-coda) coordination (in which two gestures are phased sequentially). The differential coupling modes as a function of syllable structure are hypothesized to be “part of the phonological knowledge of particular word forms” (Nam, 2007: 128). One can therefore predict that CV intergestural timing is more stable in ♯CV than C♯V, given that its in-phase coupling relationship in the former case can be considered to be lexically specified (as part of the phonological knowledge) whereas it has to be determined postlexically in the latter case (i.e., for the heterosyllabic C♯V sequence). Furthermore, actual CV coordination patterns may also differ between ♯CV and C♯V, so that the consonantal gesture in C♯V may be shifted to the left (showing less overlap with the following vowel), reflecting its underlying syllable affiliation, as compared to the consonantal gesture in ♯CV.

An alternative prediction, however, may be that there will be no syllable structure effect on CV coordination and its temporal stability (again when compared only in the Wd condition). Given that the phrase-internal Wd boundary condition is an environment in which C of the preceding syllable can be ‘resyllabified’ into the following syllable, C and V gestures may reorganize temporally as having an in-phase coupling relationship. In such a case, differences that might arise with syllable structure may be neutralized in the temporal dimension. This would also imply that the in-phase coupling mode, whether lexically or postlexically determined, would result in a commensurable phonetic effect on CV intergestural timing. In the IP boundary condition, on the other hand, the consonantal gesture in the C♯V context will be clearly shifted to the left relative to the vocalic gesture, showing less overlap with the following vowel (as compared to CV coordination in the ♯CV context). It is also expected that intergestural timing is more variable for C♯V than for ♯CV in the IP boundary condition. This is because an IP boundary is likely to block C and V gestures in the C♯V context from reorganizing temporally, thus constituting neither an in-phase nor an anti-phase coupling relationship between them, whereas C and V gestures in the ♯CV context are expected to have an in-phase coupling relationship regardless of boundary strength.

Finally, the present study considers CV coordination and its temporal stability from the perspective of the boundary effect in different syllable structure conditions: C♯V vs. ♯CV. The stability in intergestural timing can be considered in connection with the notion of *bonding strength* (i.e., the strength of intergestural cohesion) (e.g., Browman & Goldstein, 2000). The bonding strength hypothesis was originally motivated to account for temporal variability in context-sensitive intergestural timing between heterosyllabic consonant gestures. It is assumed that the larger the prosodic boundary that intervenes two gestures, the weaker the bonding strength between them, resulting in less overlap and more variability (see Cho, 2001 and Cho et al., 2007 for a related discussion). By the same logic, C and V gestures in C♯V are expected to be timed with greater variability and less overlap across an IP boundary than across a Wd boundary. For the ♯CV coordination, on the other hand, the bonding strength hypothesis becomes no longer relevant as there is no boundary between C and V gestures in ♯CV. An important question is then whether CV coordination in ♯CV also varies a function of boundary strength (IP vs. Wd), and, if so, how. CV gestures in this condition are likely to undergo some temporal expansion IP-initially (as part of domain-initial strengthening), and such temporal modification may result in greater temporal *variability* as compared to the Wd condition. But the reverse may also be possible. Given that ♯CV in the IP condition has less coarticulatory influence coming from preceding gestures across an IP boundary (e.g., Cho, 2004), there may arise greater temporal *stability* associated with IP than with Wd. Testing these possibilities will thus allow us to understand characteristics of domain-initial strengthening effects on CV in terms of CV coordination and temporal stability.

In sum, the present study systematically investigates effects of prosodic boundary and syllable structure on the temporal realization of CV gestures for /ma/ and its coordination in Korean, based on the articulatory data obtained with a magnetometer system (Electromagnetic Midsagittal Articulography (EMA), Carstens Articulograph AG 200). Answers to the questions addressed above will illuminate the dynamical nature of CV timing in speech production with some theoretical implications for the temporal scope of the boundary effect, the gestural coupling (that determines CV intergestural timing) and the π-gesture (that modulates the timing, reflecting the boundary strength).

Table 1
Test speech materials. The angled brackets ‘[ ]’ are used to indicate an intended prosodic grouping of speech materials into intonational phrases, and ‘♯’ refers to a prosodic boundary.

Conditions		Target-bearing sentences
♯=Wd	♯CV (onset) /i#ma/	[jʌŋlani ʌmʌni **# ma**nɨl] [tʃalk*atʃitʌla] ‘mother's ♯ garlic’ *Young-Ran's mother's garlic peels easily.* Korean: [영란이 어머니#마늘] 잘 까지더라..
	C♯V (coda) /im#a/	[jʌŋlani ʌmʌni**m # a**tɨl] [tʃaŋkakas*tʌla] ‘mother's ♯ son’ *Young-Ran's mother's son has married.* Korean: [영란이 어머님#아들] 장가갔더라..
♯=IP	♯CV (onset) /i#ma/	[jʌŋlani ʌmʌni] ♯ [matʃʰim tʃal osjʌs*ʌjo] ‘mother’ ♯ ‘just in time’ *Young-Ran's mother, you came just in time.* Korean: [영란이 어머니,] # [마침 잘 오셨어요.]
	C♯V (coda) /im#a/	[jʌŋlani ʌmʌnim] ♯ [atʃʰim tʃalmʌkʌs*ʌjo] ‘mother’ ♯ ‘breakfast’ *Young-Ran's mother, I had such a good breakfast.* Korean: [영란이 어머님,] # [아침 잘 먹었어요.]

2. Methods

2.1. Participants and speech materials

Four Seoul Korean speakers (three male and one female) who were in their 20s, born and raised in Seoul participated in this experiment. They were not aware of the purpose of the experiment; and they were paid for the participation after the experimental session.

To examine effects of boundary and syllable structure on temporal realization of CV gestures in the articulatory space, the /ima/ target segment sequences were created with /i#ma/ and /im#a/, which occurred in meaningful carrier sentences as shown in Table 1. The consonant was always a bilabial /m/ (which is generally known to interfere minimally with the tongue body movement) and the (following) vowel was /a/ in the /i/-to-/a/ context. The consonant was located either in coda or in onset position (i.e., C♯V or ♯CV, where ♯ = a prosodic boundary). In the coda condition (C♯V), C and V were underlyingly heterosyllabic belonging to different lexical items, and in the onset condition (♯CV), they were underlyingly tautosyllabic belonging to the same lexical item. The prosodic boundary within the target sequences was varied from a Word (Wd) boundary to an Intonational Phrase (IP) boundary.[2] An IP may be produced with different degrees of boundary strength, such that boundary strength, for example, is likely greater with than without a pause (e.g., Krivokapić & Byrd, 2012). But the present study did not include tokens with a pause, maintaining the IP-induced boundary strength more or less the same (see below for further explanation).[3]

2.2. Procedure

The articulatory data acquisition took place at the Hanyang Phonetics and Psycholinguistics Laboratory (HPPL). The 2D Electromagnetic Midsagittal Articulography (Carstens AG200) was used to track articulatory movements (see Tuller, Shao, & Kelso, 1990 and Hoole, 1996, for more technical information on the Carstens system). Seven sensor coils were used with five sensors attached on articulators: the tongue body, the tongue tip, the upper and lower lips at the vermilion borders, and one at the lower gumline of the mandibular incisor for monitoring the jaw movement. (For the purpose of the present study, the data obtained from the tongue body and the lips were analyzed; see Section 2.3 for more detail.) Additional two pellets were used as reference points, attached on the upper gumline and the nose bridge to correct for the head movement inside the helmet. The exact location of the sensor coils on the tongue body varied from speaker to speaker, depending on the size of the tongue, but it was placed on the rearmost point when the tongue was pulled out, which was about 4.5–5.5 cm from the tongue tip. In addition, a bite plate with two extra sensors was used to rotate the obtained data, so that the occlusal plane obtained by the two coordinate points created on the bite plate became the horizontal (x) axis of the data, and the y axis became perpendicular to the occlusal plane. Entire articulatory movement data were sampled at 200 Hz and low-pass filtered at a cut-off frequency of 20 Hz. All the filtering and rotation processes were performed by the TAILOR program (Carstens' data processing program).

Target-bearing sentences in Table 1 were repeated about 15 times per speaker in a pseudo-randomized order. The prosodic boundary information of each utterance token was examined by all three authors who were trained Korean ToBI transcribers. All the utterances were also visually inspected, with displays of the spectrograms and acoustic waveforms. Only those tokens whose boundary information was agreed upon by all three authors were included for further analyses. Speakers always produced an IP boundary with the sentences in Table 1 that were expected to elicit an IP boundary along with a major syntactic juncture. For the Wd tokens, speakers produced the intended phrase-internal Word boundary juncture about 85.8% of the time (103 out of 120 Wd tokens). Furthermore, some utterance tokens which had irregular velocity profiles in the movement data were excluded. In order to have a comparable number of tokens across speakers, we aimed to include 10 repetitions which were randomly selected from the data to be analyzed in the present study. As a result of the prosodic transcription and the data processing, however, we ended up with repetitions ranging from 8 to 10 per condition.

2.3. Measurements

For the purpose of the present study, we defined articulatory landmarks using a Matlab-based kinematics analysis software, MVIEW (developed at Haskins laboratory). For the consonantal gesture, we used the kinematic data obtained by the Lip Aperture (the Euclidean distance between the upper and the lower lips). Fig. 1 shows CV articulatory landmarks along with CV duration and intergestural timing measures. The onset and the target attainment points of the lip closing movement and the tongue lowering movement (in the vertical tongue movement dimension) were identified from a threshold-crossing point in the velocity profile. The threshold was set to 10% of the range from the minimum and the maximum local movement velocities. Ideally, movement onsets and targets should be defined as zero-crossing points in the velocity profile, but due to the inherent noise in the articulatory movement data, a threshold was used to minimize it.

At this point it is worth discussing possible measurement difficulties in detecting the onset of the /i/-to-/a/ tongue movement for the vowel especially at an IP boundary. When an IP boundary is produced with a substantial pause, the tongue movement gesture from /i/-to-/a/ is expected to be delayed during the pause (as there is no gestural activation for the following /a/ yet). During the pause the already raised tongue body for the preceding vowel /i/ may then be gradually lowered toward its neutral position. When the following gesture for /a/ is subsequently activated, the tongue is expected to continue to go down to reach the target for /a/. This would make it hard to determine the exact onset point of the tongue lowering movement from /i/ to /a/. Fig. 2 illustrates two cases, one with a clear tongue movement onset with no pause (Fig. 2a) and one with an obscured tongue movement onset with pause (Fig. 2b). In the preliminary articulatory data (which were collected from four speakers), we encountered so many such obscured cases, which led us to abandon the entire data set and acquire an additional data set. To obtain the new data set with no substantial pause (which was used for the present study), speakers were asked to produce IP sentences connectedly without any substantial break during the utterance. In some occasions speakers were

[2] Korean also has a smaller phrase called the Accentual Phrase (AP) which is characterized primarily by a particular tonal pattern (e.g., LHLH) (Jun, 1998). While it would be interesting to see how the temporal organization may be further modified across a larger versus a smaller phrase boundary, the present study focused on a phrase boundary versus a word boundary leaving this question open for future research.

[3] It is worth noting that syntactic categories of the IP-initial words were different in the onset versus the coda conditions. The /m/-initial word [matʃʰim] ('just in time') was used as an adverb whereas the /a/-initial word [atʃʰim] ('breakfast') was used as a subject noun. What was critical for the purpose of the present study, however, was to induce an IP boundary and match the segmental contexts as closely as possible except for the syllable structure. One cannot entirely preclude a possibility that syntactic categories influence fine-grained phonetic realizations, especially given that the syntactic structure may influence its prosodic phrasing (Cooper & Paccia-Cooper, 1980; Nespor & Vogel, 1986; Selkirk, 1986). We, however, assume that once a prosodic structure has been created for a given sentence, segmental realization of its constituents at the phonetic level is modulated directly by the prosodic structure, independently of syntactic structure (see Keating & Shattuck-Hufnagel, 2001, for a related discussion).

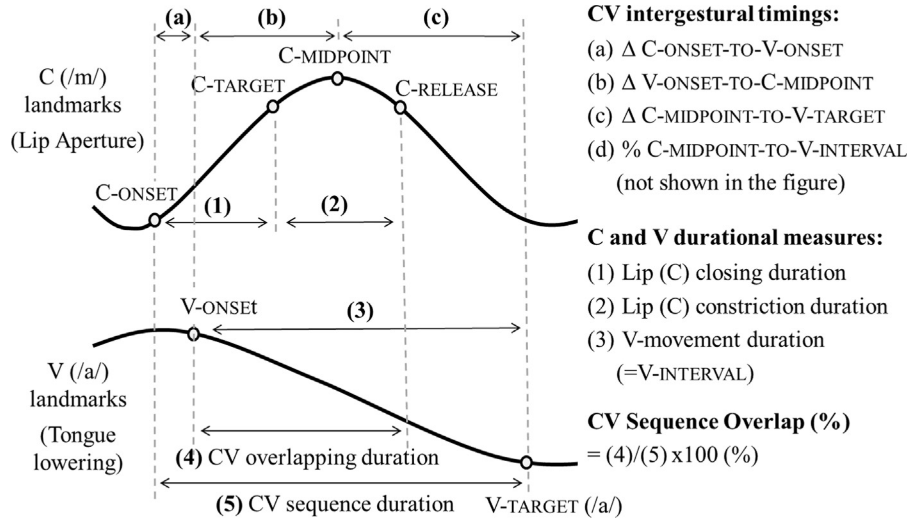

Fig. 1. Schematized trajectories of lip (closing and opening) movement and tongue (lowering) movement with an indication of the measured temporal variables.

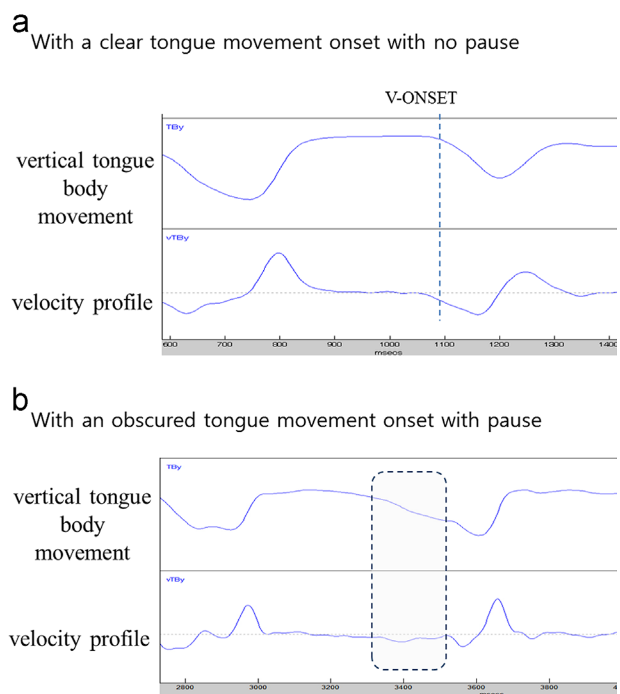

Fig. 2. Representative tongue movement profiles with a clear and an obscured /i/-to-/a/ movement onset. The dotted box in (b) indicates a hypothetical period of an articulatory pause during which the tongue may gradually go down towards its neutral position until the following gesture is activated.

asked to repeat the token when a pause was noticed by the experimenter. The obtained data were then further inspected both auditorily and visually (with displays of the spectrogram and acoustic waveforms) by all three authors, and about 18.4% of the tokens (22 out of 120 IP tokens) that were agreed to have a pause were excluded.

To examine how C and V durations would vary as a function of boundary and syllable structure, three durational measures were included (see Fig. 1 for visualization of these measures):

(1) *C-closing duration (ms)*: the interval from C-onset to C-Target.
(2) *C-closure (constriction) duration (ms)*: the interval from C-target to C-release.
(3) *V-movement duration (ms)*: the interval from V-onset to V-target.

To examine CV intergestural timings and temporal overlap, the following five measures were included (see also Fig. 1 for visualization of these measures):

(a) *ΔC-onset-to-V-onset (ms)*: the timing between the lip closing movement onset and the tongue lowering movement onset;
(b) *ΔV-onset-to-C-midpoint (ms)*: the timing between the tongue lowering movement onset and the *C* constriction midpoint;
(c) *ΔC-midpoint-to-V-target (ms)*: the timing between the midpoint of lip constriction to the target of the tongue lowering movement;
(d) *%C-midpoint-to-V-interval*: the relative timing of C-midpoint during the vowel interval (measure (3) in Fig. 1) on a normalized (0–100) scale; and

(e) *CV Sequence Overlap (%)*: a percent of overlapping duration (the measure (4) in Fig. 1) relative to the total CV sequence duration (measure (5) in Fig. 1).

Note that the CV timing measure (a) Δ*C-onset-to-V-onset* was included in order to examine the extent to which onsets of C and V movements are synchronized; the measures, (b) Δ*V-onset-to-C-midpoint*, and (c) Δ*C-midpoint-to-V-target* were based on the timing of C constriction midpoint relative to the V-movement onset and the V-movement target. C constriction midpoint (often known as C-Center) was used as the consonantal anchor point as it is known as one of the important consonantal landmarks that are coordinated with the vocalic gesture even for the singleton consonant (e.g., Nam, 2007). While the measures (a)–(c) were based on absolute timing values, the measure (d) *%C-midpoint-to-V-interval* was based on a relative timing of *C* constriction midpoint during the vowel interval, expressed on a 0–100 scale, so that 0% means that C-midpoint is aligned with the left edge of V-interval (the V-movement onset) and 100%, with the right edge of V-interval (the V-movement target). Finally, the measure (e) *CV Sequence Overlap* (%) was used to examine how much C and V are temporally overlapped. Finally, CV intergestural timing stability was examined by comparing both (absolute) standard deviations and relative standard deviations (coefficients of variation) of three CV timing measures: the measures (b)–(d) that involved the timing of C-midpoint relative to V-onset and V-target in absolute and relative terms (i.e., Δ*V-onset-to-C-midpoint*, Δ*C-midpoint-to-V-target* and *%C-midpoint-to-V-interval*).

It should be noted that two absolute CV timing measures that involved C-midpoint (i.e., Δ*V-onset-to-C-midpoint* and Δ*C-midpoint-to-V-target*) revealed CV timing patterns comparable to those found with the relative CV timing (*%C-midpoint-to-V-interval*).[4] So for the sake of simplicity, only the findings with the relative CV timing measure will be reported in full in the results section. (Readers are referred to Appendix A for a full report on the results with the absolute measures.)

2.4. Statistical analysis

The effects of prosodic boundary and syllable structure on various kinematic measures were statistically evaluated by conducting a series of univariate Analyses of Variance (ANOVAs), using SPSS 18.0 for Windows. Boundary (IP vs. Wd) and Syllable structure (in C#V vs. in #CV) were employed as fixed factors and Speaker as a random factor. With a random factor, SPSS employs the Satterthwaite approximation to calculate denominator degrees of freedom in a conservative way with no assumptions of equal variances (see Marin, 2013, for a related discussion on using this method in analyzing similar articulatory data.) (It is worth noting that the results of the univariate ANOVAs with Speaker as a random factor were largely comparable to those of repeated measures ANOVAs with each speaker contributing one averaged score per condition, as both methods use the same (reduced) denominator degrees of freedom.) Posthoc tests were performed by running one-way ANOVAs (again with Speaker as a random factor) along with running eta(η) statistics which were used to estimate the effect size of mean differences. To examine the intergestural timing stability in three measures (i.e., Δ*V-onset-to-C-midpoint*, Δ*C-midpoint-to-V-target* and *%C-midpoint-to-V-interval*), standard deviations of means were qualitatively compared between conditions, along with results of ANOVAs run on standard deviations and coefficients of variation (relative standard deviations). Statistical results with a *p*-value less than 0.05 were considered significant, and results with a *p*-value less than 0.1 were noted as trend effects.

3. Results

3.1. CV durational measures

3.1.1. Lip closing movement duration and constriction duration for C

A univariate ANOVA returned a significant main effect of Boundary on Lip closing (movement) duration—i.e., it was longer in the IP than in the Wd condition (mean diff. 25.3 ms; $F[1,3]=11.15$, $p<0.05$). Syllable structure, on the other hand, yielded a trend effect, showing a tendency towards a longer Lip closing duration in the coda (C#V) condition than in the onset (#CV) condition (mean diff., 6.7 ms, $F[1,3]=7.09$, $p<0.08$). Importantly, there was an interaction between boundary and syllable structure, though at a trend level ($F[1,3]=5.88$, $p<0.1$). Results of posthoc tests indicated that the interaction, as can be seen in Fig. 3a, was partially because the syllable structure effect was observable only in the IP condition in which Lip closing duration tended to be longer in coda (C#V) than in onset (#CV) position (mean diff., 12.1 ms, $F[1,3]=6.50$, $p<0.09$, $\eta^2=0.68$), while no coda-onset difference was observed in the Wd boundary condition (mean diff., 1.2 ms, $F[1,3]=1.84$, $p>0.1$, $\eta^2=0.38$). The interaction was also attributable to the fact that the size of the boundary effect differed as a function of syllable structure—i.e., the IP-induced lengthening of Lip closing duration was more robust when the consonant occurred in domain-final coda position (C̲#V) (IP vs. Wd mean diff., 31 ms, $F[1,3]=149.5$, $p<0.01$, $\eta^2=0.98$) than when it occurred in domain-initial onset position (#C̲V) (IP vs. Wd mean diff., 19 ms, $F[1,3]=7.49$, $p<0.08$, $\eta^2=0.71$).

Lip constriction (closure) duration showed similar results as Lip closing duration (Fig. 3b). There was a significant main effect of Boundary on Lip constriction duration, which was longer for IP than for Wd (mean diff., 17.1 ms, $F[1,3]=25.21$, $p<0.05$). The Syllable structure factor yielded a trend effect, showing a tendency towards a longer Lip constriction duration in C̲#V than in #C̲V (mean diff., 3.0 ms; $F[1,3]=7.55$, $p<0.07$), which also interacted with Boundary as a trend effect ($F[1,3]=6.2$, $p<0.09$). The interaction was in part because Lip constriction duration tended to be longer in coda than in onset position when the comparison was made in the IP boundary condition (mean diff., 7.0 ms, $F[1,3]=6.68$, $p<0.08$, $\eta^2=0.69$), while no coda-onset difference was observed in the Wd boundary condition (mean diff., 1.3 ms, $F[1,3]=2.59$, $p>0.1$, $\eta^2=0.46$). The interaction effect on Lip constriction duration also stemmed from the fact that the boundary effect was more robust on the consonant in (domain-final) coda than in (domain-initial) onset position (coda: IP vs. Wd mean diff., 21 ms, $F[1,3]=125.8$, $p<0.01$, $\eta^2=0.98$; onset: IP vs. Wd mean diff., 13 ms, $F[1,3]=13.59$, $p<0.05$, $\eta^2=0.82$).

[4] We also measured the timing between V-onset and C-target (closing target) and C-release to V-target, but decided not to report the results in the paper as the patterns found with these timing measures were very similar to those found with the timing involving the constriction midpoint.

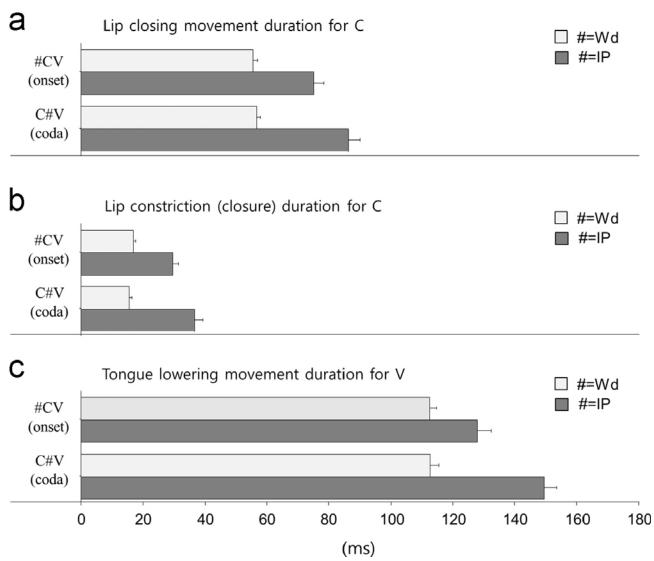

Fig. 3. Boundary × syllable structure interaction effects on three temporal measures: (a) lip closing movement duration, (b) lip constriction duration, and (c) tongue lowering movement duration. Error bars refer to standard errors.

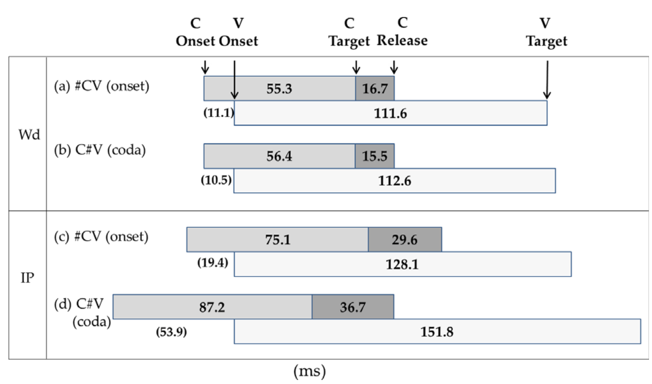

Fig. 4. An overview of CV coordination and CV overlap as a function of interaction between boundary and syllable structure. Dark gray bars correspond to lip constriction duration; light gray bars to lip closing movement duration; and white bars to V-movement duration.

3.1.2. Tongue lowering movement duration (V-movement duration)

There was a significant main effect of boundary and syllable structure on V-movement (tongue lowering) duration. It was longer for IP than for Wd (Boundary effect, mean diff., 27.8 ms, $F[1,3]=12.58$, $p<0.05$); and it was longer in the V-initial (C♯V) condition than in the C-initial (♯CV) condition (Syllable structure effect, mean diff., 12.4 ms, $F[1,3]=121.07$, $p<0.001$). Again, there was a significant Boundary × Syllable structure interaction ($F[1,3]=34.2$, $p<0.01$). As shown in Fig. 3c, the interaction was due to the fact that the boundary-induced lengthening of V-movement duration (the Boundary effect) was significant only in the V-initial (coda, C♯V) condition (IP vs. Wd mean diff., 39.1 ms, $F[1,3]=18.24$, $p<0.05$, $\eta^2=0.86$), while the boundary effect was marginal in ♯CV (IP vs. Wd mean diff., 16.4, $F[1,3]=5.56$, $p<0.1$, $\eta^2=0.64$). That is, the boundary-induced temporal expansion of V-movement duration was reliable only when V was strictly domain-initial in C♯V. The interaction was also accountable by an asymmetric Syllable structure effect as a function of boundary type. The Syllable structure effect (i.e., the lengthening of V-movement duration in the V-initial (C♯V) vs. the C-initial (♯CV) condition) turned out to be significant only in the IP condition (mean diff., 23.6 ms, $F[1,3]=32.6$, $p<0.05$, $\eta^2=0.91$; compare dark gray bars in Fig. 3c), but not in the phrase-internal, Wd boundary, condition (mean diff. 1.1 ms, $F[1,3]<1$, $p>0.1$, $\eta^2=0.08$; compare light gray bars in Fig. 3c).

3.2. CV intergestural timing and overlap

Overall patterns of CV intergestural timing and overlap are illustrated in Fig. 4. A quick qualitative look at the figure appears sufficient to appreciate that CV intergestural timings differ between boundary and syllable structure conditions in some meaningful ways. One notable difference is that CV timing as depicted in Fig. 4d (i.e., when CV belongs to different syllables, separated by an IP boundary) is clearly distinct from that in any of the other three conditions in Fig. 4a–c. In the following sections, detailed statistical results on CV timing are reported with reference to the patterns shown in Fig. 4.

3.2.1. CV timing with respect to C and V movement onsets

CV timing between C-onset (lip closing movement onset) and V-onset (tongue movement onset) can be seen in Fig. 4 in which the data are aligned to V-onset. Even at a glance it is obvious that C-onset starts substantially earlier relative to V-onset especially in one particular condition—i.e., in C♯V with an IP boundary (Fig. 4d). That is, when C and V were separated by an IP boundary, C-onset leads V-onset by some 50 ms, whereas the lead of

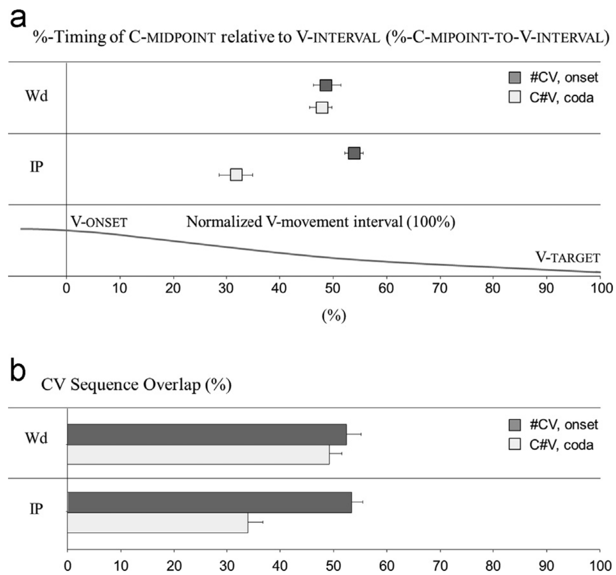

Fig. 5. Boundary × syllable structure interaction effects on CV timing and overlap measures: (a) %-C-MIDPOINT-TO-V-INTERVAL and (b) CV Sequence Overlap (%).

Table 2
Standard deviations (SDs), relative standard deviations (coefficients of variation, RSDs) and means of CV timing measures for each condition pooled across speakers ('⁎' refers $p<0.05$; and the largest and the smallest SD and RSD values are in bold).

CV timing measures		C♯V (coda) IP	C♯V (coda) Wd	♯CV (onset) IP	♯CV (onset) Wd
(a) ΔV-ONSET-to-C-MIDPOINT	SD	**17.7**	12.7	15.6	**11.6**
	RSD	**0.41**	0.29	**0.19**	0.29
	(Mean)	(49.2)	(54.2)	(70.5)	(54.8)
SD: Boundary, $F[1,3]<1^{n.s.}$; Syllable, $F[1,3]=1.39^{n.s.}$; Boundary × Syllable, $F[1,3]<1^{n.s.}$					
RSD: Boundary, $F[1,3]<1^{n.s.}$; Syllable, $F[1,3]<1^{n.s.}$; Boundary × Syllable, $F[1,3]<1^{n.}$					
(b) ΔC-MIDPOINT-to-V-TARGET	SD	**23.0**	10.4	**5.3**	9.8
	RSD	**0.23**	0.18	**0.09**	0.16
	(Mean)	(100.2)	(58.7)	(57.6)	(57.8)
SD: Boundary, $F[1,3]=8.17^{tr.}$; Syllable, $F[1,3]=18.12_{*}$; Boundary × Syllable, $F[1,3]=23.47_{*}$					
RSD: Boundary, $F[1,3]<1^{n.s.}$; Syllable, $F[1,3]=7.07^{n.s.}$; Boundary × Syllable, $F[1,3]=5.22^{tr.}$					
(c) %-C-MIDPOINT-TO V-INTERVAL	SD	**10.6**	8.3	**4.9**	8.3
	RSD	**0.24**	0.21	**0.11**	0.20
	(Mean)	(31.8)	(47.5)	(53.4)	(48.4)
SD: Boundary: $F[1,3]<1^{n.s.}$; Syllable: $F[1,3]=16.90_{*}$; Boundary × Syllable: $F[1,3]=21.42_{*}$					
RSD: Boundary, $F[1,3]<1^{n.s.}$; Syllable, $F[1,3]=2.51^{n.s.}$; Boundary × Syllable, $F[1,3]=5.89^{tr.}$					

C-ONSET in the other three conditions was relatively small, falling in the range of 10.5–19.4 ms. Results of a series of *t*-tests indicated that the observed earlier timing of C-ONSET (i.e., leftward shifting of C relative to V) was significant only when C and V were separated by an IP boundary (C♯V where ♯=IP, $t(3)=4.31$, $p<0.05$), while the numeric leads in the other conditions were statistically negligible (all at $p>0.1$).

3.2.2. Timings of C-MIDPOINT during the vowel

Fig. 4 also shows differences in CV timing in terms of timing of C constriction during the vowel. Again at a glance, an interaction between boundary and syllable structure can be observed: when CV timings are compared between C♯V and ♯CV just in the IP condition (Fig. 4c vs. d), C constriction (dark gray bars) tends to be shifted to the left in the C♯V (coda) condition (getting farther away from V-TARGET, the right edge of V-movement), but to the right in the ♯CV (onset) condition (getting closer to V-TARGET). No such shifting due to syllable structure, however, was observed in the Wd boundary condition (Fig. 4a vs. b).

The qualitatively observed Boundary by Syllable Structure interaction effects were further confirmed by results of an ANOVA which was run on the relative measure %-C-MIDPOINT-TO-V-INTERVAL: the relative timing data of C midpoint expressed as a percentage of its timing relative to V-movement interval. (See Appendix A for similar results with absolute timing measures ΔV-ONSET-TO-C-MIDPOINT and ΔC-MIDPOINT-TO-V-TARGET.) Crucially, there was a Boundary by Syllable Structure interaction ($F[1,3]=12.25$, $p<0.05$), as illustrated in Fig. 5a. (There was no main effect of Boundary ($F[1,3]<1$, $p>0.1$), while Syllable structure returned a significant main effect ($F[1,3]=56.67$, $p<0.01$).) Results of posthoc tests on this relative measure indicated that the Syllable structure effect (i.e., C midpoint came earlier in C♯V than in ♯CV) was significant only in the IP boundary condition (32.6% vs. 53.4%, $F[1,3]=26.4$, $p<0.05$, $\eta^2=0.9$), showing a leftward shifting of C in the IP/coda (C♯V) condition. No difference due to Syllable structure was observed in the Wd boundary condition (47.1% vs. 47.4%, $F[1,3]<1$, $p>0.1$, $\eta^2<0.01$).

From the perspective of the boundary effect, the Boundary by Syllable Structure interaction was due to an asymmetric boundary effect as a function of syllable structure. There was a trend effect of boundary on %-C-MIDPOINT-TO-V-INTERVAL only when the boundary conditions were compared in the C♯V condition—i.e., C midpoint in C♯V tended to come earlier during the vowel across an IP than a Wd boundary (32.6% vs. 47.1%, $F[1,3]=7.23$,

$p<0.07$, $\eta^2=0.69$). However, no such boundary effect was observed in ♯CV (53.4% vs. 47.4%, $F[1,3]<1$, $p>0.1$, $\eta^2=0.22$), indicating that boundary strength did not bring about modification of CV timing in domain-initial position.

3.2.3. Sequence Overlap

Sequence Overlap (%) is a percent of overlapping duration between C duration (the closing movement duration plus the constriction duration) and V-movement duration compared to the total CV sequence duration. A univariate ANOVA run on the Sequence Overlap data returned a significant main effect of Syllable Structure ($F[1,3]=27.98$, $p<0.05$), but no effect of Boundary ($F[1,3]=1.4$, $p>0.1$). As was the case with other CV timing measures there was a significant Boundary by Syllable Structure interaction ($F[1,3]=13.45$, $p<0.05$), which is shown in Fig. 5b. Results of posthoc tests indicated that the amount of CV Sequence Overlap (%) was significantly smaller in the C♯V condition across an IP boundary than that in any of the other three conditions ($p<0.05$). The smallest amount of CV Sequence Overlap (%) in C♯V across an IP boundary indicates that C midpoint was shifted to the left relative to the V-movement interval in this condition. It is again worth noting that the Syllable structure effect (less overlapping for C♯V than for ♯CV) was significant only in the IP condition (C♯V vs. ♯CV, mean diff., −18.6%, $F[1,3]=32.3$, $p<0.05$, $\eta^2=0.92$) while the syllable effect disappeared in the Wd boundary condition, showing neutralization in CV overlap between C♯V and ♯CV (C♯V vs. ♯CV, mean diff., −2.5%, $F[1,3]<1$, $p>0.1$, $\eta^2=0.24$). Likewise, the Boundary effect (less overlapping for IP than for Wd) was found to be significant only in C♯V (IP vs. Wd, mean diff., −14%, $F[1,3]=9.91$, $p<0.05$, $\eta^2=0.92$), while the boundary strength did not yield any significant difference in CV Sequence Overlap in ♯CV (IP vs. Wd, mean diff., 2.1%, $F[1,3]<1$, $p>0.1$, $\eta^2=0.05$), which again confirmed that CV timing in domain-initial position was not modified by boundary strength.

3.3. Stability of CV intergestural timing

Stability of CV timing was examined by comparing standard deviations (SDs) of mean values and coefficients of variation (RSDs, relative standard deviations) for three interrelated CV timing measures: the interval from V onset to C midpoint (i.e., ΔV-ONSET-TO-C-MIDPOINT), the interval from C midpoint to V target (i.e., ΔC-MIDPOINT-TO-V-TARGET) and percent timing of C midpoint relative to V-movement interval (i.e., %-C-MIDPOINT-TO-V-INTERVAL). Table 2 shows SDs, RSDs (coefficients of variation) and means of these measures for each condition pooled across speakers. Note that in the figure the smallest and the largest SDs and RSDs in each condition are printed in bold, along with a summary of results of ANOVAs run on SD and RSD values.

For all three CV timing measures, the largest SDs and RSDs were both found consistently in C♯V with an IP boundary (see values in bold in the first numeric column in Table 2), suggesting that CV timing was least stable in C♯V separated by an IP boundary. In other words, a larger boundary induced a greater temporal variability when *C* and *V* belonged to different lexical items in C♯V. The smallest SD and RSD values, on the other hand, were observed consistently in ♯CV, showing that CV timing was more stable when they belonged to the same lexical item (in ♯CV) than when they belong to different lexical items (in C♯V). In the ♯CV condition, all three measures showed smaller RSD values for IP than for Wd, suggesting that CV timing was more stable in an IP than in a Wd boundary condition. For SD values, the latter two timing measures (ΔC-MIDPOINT-TO-V-TARGET and %-C-MIDPOINT-TO-V-INTERVAL) showed smaller SD values for IP and for Wd. Thus the distribution of both SD and RSD across conditions indicates that the most stable CV timing pattern is associated with ♯CV gestures at an IP boundary condition.

Statistically, the latter two timing measures (ΔC-MIDPOINT-TO-V-TARGET and %-C-MIDPOINT-TO-V-INTERVAL) showed Boundary by Syllable Structure interactions, while the first measure (ΔV-ONSET-TO-C-MIDPOINT) did not show any statistically significant differences (see the statistical summaries in Table 2). Results of posthoc tests run on the latter two CV timing measures were largely in support of the qualitative observations on the distribution of standard deviations and coefficients of variation (RSD) in Table 2. For ΔC-MIDPOINT-TO-V-TARGET, the SD for IP in C♯V was significantly larger than that for Wd in the same C♯V condition (23.0 vs. 10.4, $F[1,3]=41.96$, $p<0.01$, $\eta^2=0.93$), although the RSD showed only a trend effect (0.23 vs. 0.19, $F[1,3]=5.13$, $p<0.1$, $\eta^2=0.45$). For %-C-MIDPOINT-TO-V-INTERVAL the same pattern was observed as a trend effect for the SD (10.6 vs. 8.3, $F[1,3]=6.99$, $p<0.07$; $\eta^2=0.68$) while the difference in RSD did not reach significance (0.24 vs. 0.21, $F[1,3]<1$, $p>0.1$, $\eta^2<0.1$).

For the ♯CV condition, on the other hand, the exact opposite pattern was observed. The SD in ♯CV was significantly *smaller* for IP than for Wd for the absolute measure ΔC-MIDPOINT-TO-V-TARGET (SD: 5.3 vs. 9.8, $F[1,3]=8.67$, $p<0.05$, $\eta^2=0.82$) and the same pattern was observed as a trend effect with the RSD (0.09 vs. 0.16, $F[1,3]=6.21$, $p<0.09$, $\eta^2=0.52$). The relative measure %-C-MIDPOINT-TO-V-INTERVAL also showed the same tendency towards smaller SD and RSD for IP than for Wd in ♯CV (SD: 4.9 vs. 10.6, $F[1,3]=6.21$, $p<0.1$; $\eta^2=0.64$; and RSD: 0.11 vs. 0.20, $F[1,3]=5.97$, $p<0.1$; $\eta^2=0.37$). It is also noteworthy that while there was the opposite boundary effects on C♯V vs. ♯CV intergestural timings, there was no syllable structure effect (no difference between C♯V and ♯CV) in the Wd boundary condition in line with the temporal neutralization observed with other CV timing measures.

4. General discussion

4.1. Asymmetric boundary-related lengthening effects in C♯V vs. ♯CV

Results regarding lip closing movement duration and lip constriction duration for /m/ showed boundary-induced lengthening regardless of whether the consonant was in coda or in onset position (i.e., before (C♯V) or after (♯CV) a boundary). The size of the boundary effect, however, was larger on the preboundary than on the postboundary consonant, showing an asymmetric boundary-induced lengthening effect on the consonant. Results also revealed an asymmetric boundary-induced lengthening effect on V-movement (tongue lowering) duration: It was longer IP-initially than Wd-initially only when the vowel was strictly domain-initial (in C♯V̲), while the effect in ♯CV̲ was only marginally significant. At the outset of the paper, we questioned whether the weak or null boundary effect on the vowel in ♯CV often observed in the acoustic dimension would be because the boundary effect on the actual articulatory vowel duration is indeed attenuated after an onset consonant or because the domain-initial lengthening of the consonant is articulatorily extended into the vowel, masking the boundary effect on the vowel in the acoustic dimension. The asymmetric boundary effect on the vowel suggests that the not-strictly-initial vowel (♯CV̲) is indeed attenuated temporally in the articulatory dimension, compared to the strictly domain-initial vowel (C♯V̲), and that the frequently observed lack of the boundary effect on the acoustic vowel duration is not simply due to consonantal lengthening that might encroach the acoustic vowel duration.

The asymmetry further implies that the scope of domain-initial lengthening (as part of domain-initial strengthening) is indeed largely restricted to the initial segment whether consonantal or vocalic (e.g., Fougeron, 2001). However, the fact that the non-initial vowel (♯CV̲) also shows some degree of lengthening (though as a trend effect) implies that the effect is not necessarily blocked by the first segment, but rather it can permeate, though to an attenuated degree, into the following vowel along the articulatory dimension. This is in line with the view that the locality hypothesis is better characterized as a gradient constraint rather than as an all-or-none constraint (e.g., Byrd & Saltzman, 2003; Byrd et al., 2006; Cho, 2005; Cho & Keating, 2009; Cho & McQueen, 2005).

These findings, however, are contradictory to some other predictions that can be made in the theoretical framework of Articulatory Phonology (e.g., Browman & Goldstein, 1990, 1992; Goldstein & Fowler, 2003) and the π-gesture model (Byrd & Saltzman, 2003; Byrd et al., 2006). Given that CV gestures overlap in time with a possible synchrony of onsets (in an in-phase coupling mode) (e.g., Browman & Goldstein, 1992, 2000; Fowler, 1983; Nam, 2007; Öhman, 1966; Perkell, 1969), and given that the vowel gestures in both ♯CV and C♯V contexts are in theory realized in a vocalic tier that is independent of a consonantal tier, the proximity of the vocalic gesture to the boundary would remain the same regardless of whether the consonant is superposed on it or not. It was therefore hypothesized, as now schematized in Fig. 6a, that the π-gesture would also have the same lengthening effect on the vowel in both ♯CV and C♯V. The result of the present study, however, showed otherwise. As schematized in Fig. 6b, the effect of the boundary on their phonetic realization turned out to be sensitive to the presence or absence of an onset consonant—i.e., the syllable structure. This also informs the π-gesture model in terms of how precisely the π-gesture can be anchored to the overlapping articulatory gestures (i.e., gestural constellations) that operate in separate functional articulatory tiers. While further work is needed to corroborate this, it appears that the anchor point of a π-gesture to an articulatory gesture needs to be adjusted according to the syllable structure and the composition of the gestural constellation that the gesture is part of.

Another important consideration pertains to the generalizability of this finding in Korean to other languages, especially to English. It has often been argued that the reason why the domain-initial lengthening effect is not manifested in the acoustic duration of vowel in English is presumably because it is reserved for phonetic manifestation of lexical stress (e.g., Barnes, 2002). In an acoustic study in Korean, Cho et al. (2011) indeed suggested a language-specific operation of boundary lengthening depending on the prosodic system of a given language—i.e. domain-initial lengthening can be more readily extended to the vowel in Korean, presumably because Korean is not restricted by the lexical prominence system. Likewise, the boundary-induced lengthening of the vocalic gesture observed in the present study may be at least partially attributable to the prosodic and dynamical systems that are unique to Korean. Further cross-linguistic articulatory studies are warranted to explore the generalizability issue in connection with the theoretical consideration of how mechanisms assumed in the π-gesture model have to be language-specifically attuned.

4.2. CV coordination and temporal stability at a Word boundary

The present study also examined how CV coordination and its temporal stability would vary as a function of whether C and V gestures belong to the same syllable (underlyingly in the same lexical item) or to different syllables (underlying across a lexical boundary). A focus was made on the potentially neutralizing phrase-internal (Wd boundary) context as this context is likely to bring about ‘resyllabification’ of the consonant originally in coda position into the following vowel. Out of the multiple temporal measures taken in the present study, however, not a single measure revealed any significant difference between ♯CV and C♯V, showing a complete neutralization in the Wd boundary context at least in the articulatory measures taken in the present study. That is, not only the temporal realization of individual C and V gestures but also their intergestural timing patterns remained the same regardless of whether C and V gestures occurred in the same lexical item (in ♯CV) or across a lexical boundary (in C♯V) in the Wd boundary condition. Regarding the CV intergestural timing, it was hypothesized to be less stable in the ‘heterolexical’ (C♯V) than in the ‘tautolexical’ (♯CV) condition even in the potentially neutralizing Wd boundary context, schematized in Fig. 7a vs. b. This was in line with the assumption in Articulatory Phonology—i.e., intergestural timing is lexically specified, and therefore it may result in more stable timing compared to postlexically determined intergestural timing. The results of the present study, however, showed that the intergestural timing in C♯V was as stable as that in ♯CV as schematized in Fig. 7c, taking support away from the hypothesis.

The lack of temporal cues to the different underlying syllabic structures between C♯V and ♯CV is in sharp contrast with the case of *enchaînement* in French. Fougeron (2007) reported that the acoustic realizations of *enchaînement* consonants (which may be potentially resyllabified with the following vowel across a Word boundary) and the flanking vowels in VC♯V were not the same as those of V♯CV with underlyingly onset consonants, and their differences were available to the listeners as cues to the underlying CV syllable structure. At the moment, an interim conclusion therefore might be that Korean and French may differ in terms of temporal organization of C♯V vs. ♯CV in the phrase-internal conditions. However, given that the present study has explored only articulatory temporal measures, it remains to be seen whether C♯V is completely neutralized with ♯CV in other fine-grained phonetic details especially in the acoustic dimension.

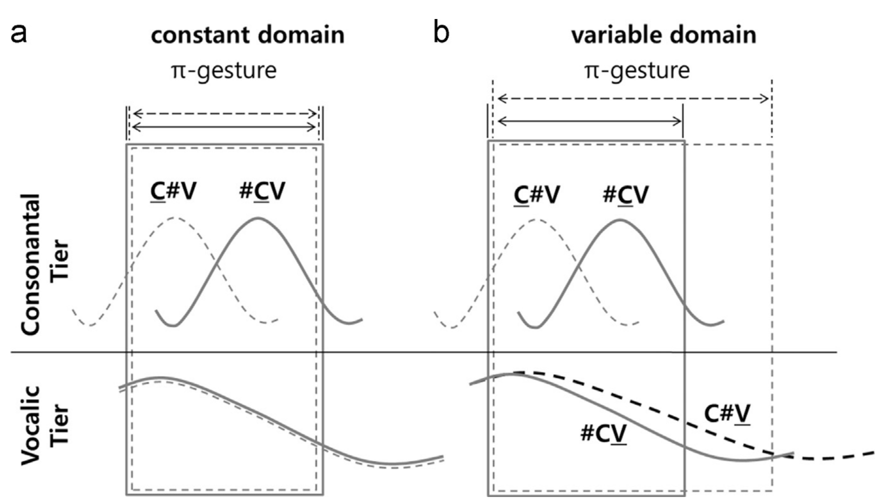

Fig. 6. Two hypothetical schematics with respect to constant vs. variable domains of the π-gesture's influence as a function of syllable structure. The solid lines are associated with ♯CV (where V is not strictly initial) and the dotted lines with C♯V (where V is strictly initial); and the arrows refer to domains of the π-gesture's influence.

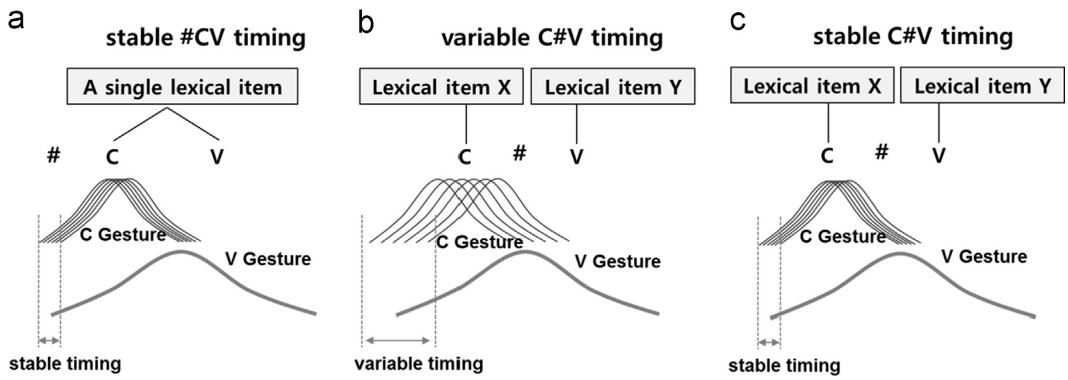

Fig. 7. Schematics of hypothesized stable vs. variable CV intergestural timings depending on whether C and V belong to the same vs. different lexical items. '♯' refers to a prosodic boundary either as a Wd or as an IP boundary, but the variable timing depicted in (b) is more likely associated with an IP boundary, and the stable timing in (c) with a Wd boundary as discussed in the text. (Note that while intergestural timings here are schematized between onsets of C and V movements, other timing measures are often considered to be more reliable such as the timing between V onset to C midpoint or between C midpoint to V target as included in the present study.).

The results regarding CV coordination and its temporal stability can be interpreted further in terms of the gestural coupling model of syllable structure (e.g., Goldstein et al., 2006; Marin & Pouplier, 2010; Mooshammer et al., 2012; Nam, 2007; Nam et al., 2009; Pouplier, 2012). While C and V gestures in ♯CV can be assumed to have a lexically determined in-phase coupling relationship (where C and V starts roughly synchronously), timing between C and V gestures in C♯V (i.e., belonging to different lexical items) has to be derived postlexically. This sparked a question as to whether a postlexically determined CV phasing relationship would be comparable to a lexically determined in-phase coupling relationship. Crucially, the results of the present study suggest that *C* and *V* gestures across a Wd boundary (in C♯V) reorganize temporally, showing a stable intergestural timing pattern (as schematized in Fig. 7c) exactly in the same way as C and V gestures are timed in a lexically determined in-phase mode (as schematized in Fig. 7a). This also implies that an in-phase coupling mode arises between C and V gestures even across a lexical boundary if it coincides with a prosodic word (Wd) boundary inside a phrase (see Mücke, Nam, Hermes, & Goldstein, 2012 for a related discussion on the postlexical modification of coupling of tone and constriction gestures in pitch accents).

At this point, it is worth addressing an issue raised in the review process with respect to whether the observed temporal neutralization between C♯V and ♯CV at a Wd boundary may be due to the articulatory dependency of the consonantal and the vocalic movements, both of which are constrained by the jaw. In the framework of Articulatory Phonology, the lip aperture and the tongue body gestures are assumed to occur on separate articulatory tiers at the (abstract) 'gestural' level. At the 'articulatory' level, however, both gestures are articulatorily implemented with a shared articulator, being the jaw. The consonantal lip opening/closing movements are influenced by the movement of the jaw, which also mediates the tongue movement due to the jaw-tongue mechanical linkage (e.g., Fletcher & Harrington, 1999; Mooshammer et al., 2007 for a related discussion). The consonantal and vocalic movements analyzed in the present study are then likely to impose conflicting demands on the jaw—i.e., the demand for the raising of the jaw linked to the lip closing movement (from /i/-to-/m/) vs. the demand for the lowering of the jaw linked to the tongue lowering movement (from /i/-to-/a/). As a result, the jaw lowering movement for the vocalic gesture is likely to be delayed (and so is the tongue lowering movement) to a similar extent in both C♯V and ♯CV conditions unless the timing is further constrained by a substantial boundary-related slowing down (as in the IP boundary condition). This may account for the similar CV timing patterns between C♯V and ♯CV observed in the Wd condition.

This possibility, however, would make much more sense if the lip closing movement and the tongue lowering movement are strictly controlled by the jaw with no degree of freedom for the lips and the tongue. But this may not be always the case. For example, the articulatory goal of lip closing can be achieved through an active lowering of the upper lip when the jaw is relatively lowered in association with a tongue dorsal gesture (see Fowler & Saltzman, 1993, for a related discussion). Furthermore, the independence of the lips from the jaw becomes clearer when the jaw is perturbed (e.g., Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984), although such a perturbed case would not represent natural speech at all. In order to resolve on this issue regarding the involvement of the jaw directly, more studies are certainly called for to investigate CV timings with consonants other than labials in different vowel sequences and with different directionality of the tongue movement.[5]

4.3. Boundary effects on CV coordination and temporal stability in C♯V vs. ♯CV

Another important goal of the present study was to explore how boundary strength would influence the way that consonantal and vocalic gestures are temporally coordinated in different syllable structures (C♯V vs. ♯CV). What has emerged from the temporal measures is that the consonant gesture is substantially shifted to the left relative to the vowel gesture, reinforcing the underlying heterosyllabicity between the two gestures across an IP boundary. This was also evident in CV Sequence Overlap (%) which showed that CV gestures were overlapped less across an IP boundary than across a Wd boundary. Furthermore, the boundary-induced reduction of CV overlap was accompanied by greater temporal variability, which was evident in significantly larger standard deviations in this heterosyllabic IP boundary condition than in any other condition. These findings can be accountable by the bonding strength hypothesis which postulates that intergestural cohesion (i.e., bonding strength) between gestures is inversely proportional to the size of the intervening prosodic boundary, so that the larger the prosodic boundary, the lesser degree of gestural cohesion and the lesser temporal stability (e.g., Browman & Goldstein, 2000; see Cho, 2001 for a related discussion). It also appears that an IP boundary blocks temporal reorganization of underlyingly heterosyllabic C and V gestures (thus having no in-phase coupling relationship). So the less stable timing

[5] To explore this issue further, we analyzed a subset of the articulatory data (i.e., only in a Wd boundary condition) with the horizontal tongue movement (from /i/-to-/a/) in the *x* axis which is likely to be influenced less by the (vertical) jaw movement. Results of a series of univariate ANOVAs (with Speaker as a random factor) run on both relative and absolute CV timing measures showed no CV timing difference between C#V and #CV conditions (%C-MID-TO-V-INTERVAL: 52.5 vs. 49.5%, $F[1,3]<1$, $p>0.1$; ΔV-ONSET-TO-C-MID: 51.9 vs. 47.7 ms, $F[1,3]=1.55$, $p>0.1$; ΔC-MID-TO-V-TARG: 44.1 vs. 44.3 ms, $F[1,3]<1$, $p>0.1$). The similar CV timings in C#V and #CV even in the horizontal tongue movement dimension indicate that the CV timings with the vertical tongue movement data presented in the results section are not entirely due to the jaw constraint.

which was predicted to occur between gestures that belong to different lexical items (as schematized in Fig. 7b) was indeed shown to occur when the lexical boundary was aligned with the IP boundary and therefore CV phasing was not governed by an in-phase coupling mode.

In the ♯CV condition, however, neither CV timing measures nor CV Sequence Overlap revealed any significant boundary effect. This indicates that CV coordination in ♯CV is not modified by boundary strength, showing invariant CV coordination, despite the fact that both C and V gestures underwent some degree of lengthening in the IP boundary (domain-initial) condition.[6] The observed invariant CV coordination is linked to the temporal stability of CV intergestural timing in ♯CV. In Section 1, we discussed a possibility that intergestural timing would be more variable at an IP than at a Wd boundary even in ♯CV because an IP boundary condition is an environment in which both C and V gestures generally undergo a substantial temporal modification. But the opposite was found to be true—i.e., the most stable intergestural timing was observed in ♯CV at an IP boundary. This could be at least in part due to the fact that ♯CV coordination has lesser contextual coarticulatory influence across an IP boundary than across a Wd boundary (e.g., Cho, 2004). (This finding is also consistent with Cole, Kim, Choi, and Hasegawa-Johnson's (2007) observation on English stops (/t,d/) which were produced with greater acoustic stability (e.g., as measured by VOT) in IP-initial than in IP-medial positions, which may be linked with articulatory precision.) Furthermore, the production of the syllable that undergoes domain-initial strengthening in the IP boundary context may resemble its canonical pronunciation form more than that of the syllable produced in the phrase-internal (Wd boundary) context. The postulated resemblance to a canonical (thus frequent) pronunciation form may explain the observed domain-initial temporal stability in ♯CV, although this speculation is subject to corroboration. Whatever other reasons there may be, however, we propose that the temporal stability in CV coordination is another characteristic of domain-initial strengthening.

5. Conclusion

The present study has investigated how prosodic boundary and syllable structure would interactively influence the temporal realization of consonantal and vocalic gestures (/m/ and /a/ in the /ima/ context), their CV coordination pattern, and temporal stability in Korean. Several important points can be drawn from the results. First, the postboundary lengthening effect was found to be largely localized to the initial segment of a prosodic constituent whether consonantal or vocalic, but the effect was taken to be gradient as the vocalic gesture also underwent some degree of lengthening in the non-initial (♯CV̲) condition. The consonant also underwent boundary-related lengthening in an asymmetric way—i.e., with more lengthening in coda than in onset position. These results were further discussed in terms of its implications for the theory of π-gesture, suggesting that the influence of π-gesture may have to be fine-tuned as a function of whether the vocalic gesture is constellated with or without an onset consonantal gesture. Second, temporal neutralization in CV timing patterns arose between the underlyingly heterosyllabic (C♯V) and the tautosyllabic (♯CV) conditions in the Wd boundary condition, indicating a possibly complete 'resyllabification' of 'C' with the following vowel in C♯V (at least in the articulatory temporal measures taken in the present study), although the phasing in C♯V has to be determined postlexically. This was interpreted as suggesting that CV gestures in the C♯V condition (belonging to different lexical items) reorganize temporally in a phrase-medial (Wd boundary) position, so that they are phased in an in-phase coupling mode in much the same way as CV gestures are in the ♯CV condition (belonging to the same lexical item). Third, in C♯V, the boundary effect was evident in the leftward shifting of the consonantal gesture relative to the vocalic gesture across an IP boundary, whereas in ♯CV, CV coordination remained invariant regardless of boundary strength, accompanied by greater temporal stability in the IP boundary (domain-initial) than the Wd boundary condition. The temporal stability associated with an IP boundary in ♯CV was proposed as an important characteristic of domain-initial strengthening.

It is hoped that these observations made in the present study have broadened our understanding of the dynamical nature of speech timing in Korean. It is also hoped that the present study gives directions to future studies in pursuing this line of research with other languages, in such a way to understand to what extent the observed temporal patterns in Korean are language-specific and are generalizable as characteristics of the human speech motor system across languages.

Acknowledgments

We thank the speakers who participated as subjects in our experiment, and the audience for their constructive comments at the workshop on *Dynamic modeling of articulation and prosodic structure* held at the University of Cologne on 7–8 May 2012. Special thanks go to the reviewers and the editor, Martine Grice, for many helpful comments and suggestions.

Appendix A. Absolute timings of C-midpoint relative to V-onset and to V-target

Results of ANOVAs run on the absolute interval data from V onset to C midpoint (ΔV-onset-to-C-Midpoint) and from C midpoint to V target (ΔC-midpoint-to-V-target) both showed a significant main effect of Syllable structure (ΔV-onset-to-C-midpoint: $F[1,3]=25.61$, $p<0.05$; ΔC-midpoint-to-V-target: $F[1,3]=359.01$, $p<0.001$), while the Boundary factor yielded a significant main effect only on ΔC-midpoint-to-V-target ($F[1,3]=14.45$, $p<0.05$). Crucially there was an interaction between syllable structure and boundary for both measures (ΔV-onset-to-C-midpoint: $F[1,3]=10.59$, $p<0.05$; ΔC-midpoint-to-V-target: $F[1,3]=12.89$, $p<0.05$), which was due to the fact that effects of Syllable structure (C♯V vs. ♯CV) were significant only in IP conditions. These interaction effects are shown in Fig. 8.

Let us first consider the interaction effect on the timing from V onset to C midpoint (ΔV-onset-to-C-midpoint). Results of posthoc tests indicated that the interaction was accountable by the Syllable structure effect being significant only in the IP boundary: ΔV-onset-to-C-midpoint was significantly longer in ♯CV than in C♯V only when the comparison was made in the IP boundary condition (mean diff., 22.5 ms, $F[1,3]=16.16$, $p<0.05$, $\eta^2=0.84$). This can be seen from the left end (V onset) to the C midpoint in Fig. 8 (in which the dotted vertical line refers to C midpoint). Looking at the timing of C

[6] This invariant CV coordination pattern is largely consistent with invariance of CV coordination observed in other cases in which consonantal duration varies due to intrinsic phonemic differences—e.g., singleton versus geminate in Japanese (Löfqvist, 2006); /b/ versus /p/ in English (Löfqvist & Gracco, 1999), and the three-way stop contrast (fortis, lenis, aspirated) in Korean (Son, Kim, & Cho, 2012). We are not entirely certain whether the unaltered intergestural timing pattern is a true reflex of the underlying invariance nature of CV coordination, but the results of the present study demonstrate a possible case in which invariant CV coordination is maintained despite the temporal variation of C and V gestures which, this time, is due to a higher-order prosodic boundary factor, rather than due to an intrinsic phonetic factor explored in earlier studies.

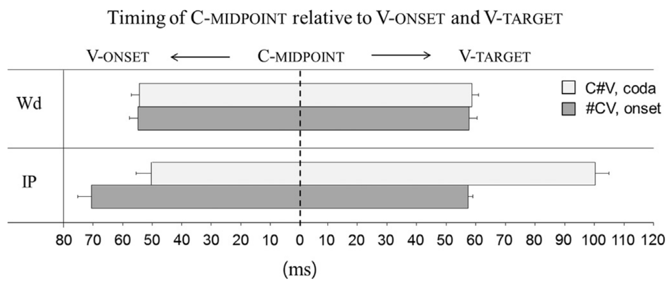

Fig. 8. Boundary × Syllable structure interaction effects on CV absolute CV timing measures: ΔV-ONSET-TO-C-MIDPOINT (left) and ΔC-MIDPOINT-TO-V-TARGET (right).

midpoint relative to the entire vowel interval (from V onset to V target), the longer interval for ΔV-ONSET-TO-C-MIDPOINT in ♯CV vs. C♯V in the IP condition (dark vs. light gray bars in the left of Fig. 8) reflects that the consonant (C midpoint) in onset (♯C̲V) position was shifted to the right during the entire V-movement interval (getting farther away from V onset) as compared to the consonant in coda (C̲♯V) position. In the Wd boundary condition, on the other hand, no such syllable structure effect was observed (mean diff., 2.1, $F[1,3]<1$, $p>0.1$, $\eta^2=0.05$), showing temporal neutralization between C♯V and ♯CV.

Similar to the interaction effect on the timing from V onset to *C* midpoint (ΔC-MIDPOINT-TO-V-TARGET), the timing between C midpoint to V target (ΔC-MIDPOINT-TO-V-TARGET) showed a mirror image of the timing between V onset to C midpoint. As shown in the right of Fig. 8, ΔC-MIDPOINT-TO-V-TARGET was significantly longer in C♯V than in ♯CV only at an IP boundary (IP: coda vs. onset, mean diff., 42.6 ms, $F[1,3]=36.64$, $p<0.01$, $\eta^2=0.92$; Wd: coda vs. onset, mean diff., 0.1 ms, $F[1,3]<1$, $p>0.1$, $\eta^2<0.01$). Again this indicates that C midpoint at an IP boundary was shifted to the left, being farther away from the right edge of V-movement interval (V-TARGET) in the coda (C̲♯V) condition compared to the onset (♯C̲V) condition, while the syllable structure effect disappeared in the Wd boundary condition, which again confirmed temporal neutralization between C♯V and ♯CV.

In sum, the two absolute measures (ΔV-ONSET-TO-C-MIDPOINT and ΔC-MIDPOINT-TO-V-TARGET) consistently showed a leftward shifting of C in coda (C♯V) position across an IP boundary, while showing temporal neutralization between C♯V and ♯CV in the Wd boundary condition. It is also worth noting that the Boundary by Syllable Structure interaction on ΔC-MIDPOINT-TO-V-TARGET also revealed that there was an effect of Boundary only in C♯V, so that the interval from C midpoint to V target (ΔC-MIDPOINT-TO-V-TARGET) was longer when C and V were separated by an IP than by a Wd boundary (mean diff. 41.1 ms, $F[1,3]=25.16$, $p<0.05$, $\eta^2=0.89$). That is, C midpoint in coda (C̲♯V) position was farther away from V-TARGET across an IP boundary than across a Wd boundary. The boundary effect, however, disappeared in ♯CV (mean diff., 1.5 ms., $F[1,3]<1$, $p>0.1$, $\eta^2=0.02$), indicating that CV timing was not modified by boundary strength when C and V belong to the same domain-initial syllable.

References

Barnes, J. A. (2002). *Positional neutralization: A phonologization approach to typological patterns (unpublished Ph.D. dissertation)*. Berkeley: University of California.
Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, *38*, 388–403.
Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.
Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.
Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and selforganization of phonological structures. *Bulletin de la Communication Parlée*, *5*, 25–34.
Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*, 209–244.
Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrase junctures. *Phonetica*, *57*, 3–16.
Byrd, D. (2006). Relating prosody and dynamic events: Commentary on the papers by Cho, Navas, and Smiljanić. In: L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory phonology 8: Varieties of phonological competence* (pp. 549–561). Berlin/New York: Mouton de Gruyter.
Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In: C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology*, Vol. 10 (pp. 31–60). Berlin/New York: Mouton de Gruyter.
Byrd, D., Kaun, A., Narayana, S., & Saltzman, E. (2000). Phase signatures in articulation. In: M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V* (pp. 70–87). Cambridge: Cambridge University Press.
Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, *120*, 1589–1599.
Byrd, D., & Saltzman, E. (2003). The elastic phase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*, 149–180.
Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, *58*, 129–162.
Cho, T. (2004). Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*, 141–176.
Cho, T. (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. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In: L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in laboratory phonology VIII* (pp. 519–548). Berlin/New York: Mouton de Gruyter.
Cho, T. (2008). Prosodic strengthening in transboundary V-to-V lingual movement in American English. *Phonetica*, *65*, 45–61.
Cho, T. (2011). Laboratory phonology. In: N. C. Kula, B. Botma, & K. Nasukawa (Eds.), *The continuum companion to phonology* (pp. 343–368). London/New York: Continuum.
Cho, T. Language effects on speech timing at the segmental and suprasegmental levels. In M. A. Redford (Ed.), *Handbook of speech production*. Wiley-Blackwell (to appear).
Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, *29*, 155–190.
Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*, 466–485.
Cho, T., & Ladefoged, P. (1999). Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*, 207–229.
Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, *39*, 344–361.
Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*, 121–157.
Cho, T., McQueen, J., & Cox, E. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243.
Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News Speech. *Journal of Phonetics*, *35*, 180–209.
Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
Edwards, J. E., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, *89*, 369–382.
Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In: W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (second ed.). Oxford: Blackwell.
Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109–135.
Fougeron, C. (2007). Word boundaries and constrast neutralization in the case of *enchaînement* in French. In: J. Cole, & J. I. Hualde (Eds.), *Papaers in laboratory phonology IX: Change in phonology* (pp. 609–642). Berlin: Mouton de Gruyter.
Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, *36*(2,3), 171–195.

Fowler, C. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General*, *112*, 386–412.
Fletcher, J., & Harrington, J. (1999). Lip and jaw coarticulation. In: W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: theory, data and technique* (pp. 144–175). Cambridge: Cambridge University Press.
Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In: M. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215–249). Cambridge: Cambridge University Press.
Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In: N. O. Schiller, & A. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). Berlin: Mouton de Gruyter.
Hawkins, S. (1992). An introduction to task dynamics. In: G. Docherty, & B. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 9–25). Cambridge: Cambridge University Press.
Hoole, P. (1996). Issues in the acquisition, processing, reduction and parameterization of articulographic data. *FIPKM (Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation der Universität München)*, *34*, 158–173.
Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*, 189–226.
Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, *97*, 157–173.
Kelso, J. A. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory coopertion following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 812–832.
Krivokapić, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, *40*, 430–442.
Löfqvist, A. (2006). Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *Journal of Acoustical Society of America*, *120*, 2872–2883.
Löfqvist, A., & Gracco, V. L. (1999). Interarticulator programming in VCV sequences: Lip and tongue movements. *Journal of the Acoustical Society of America*, *105*, 1864–1876.
Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics*, *41*, 211–227.
Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*, 380–407.
Mooshammer, T., Goldstein, L., Nam, H., & McClure, S. (2012). Bridging planning and execution: Temporal planning of syllables. *Journal of Phonetics*, *40*, 374–389.
Mooshammer, C., Hoole, P., & Geumann, A. (2007). Jaw and order. *Language and Speech*, *50*, 145–176.
Mücke, D., Nam, N., Hermes, A., & Goldstein, L. (2012). Coupling of tone and constriction gestures in pitch accents. In: P. Hoole, M. Pouplier, L. Bombien, C. h. Mooshammer, & B. Kühnert (Eds.), *Consonant clusters and structural complexity* (pp. 205–230). Berlin/New York: Mouton de Gruyter.
Nam, H. (2007). *A gestural coupling model of syllable structure (unpublished Ph.D. dissertation)*. Yale University.
Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In: F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupe (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin: Mouton de Gruyter.
Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of Acoustical Society of America*, *39*, 151–168.
Perkell, J. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. *Research Monograph*, vol. 53. Cambridge: The M.I.T. Press.
Pouplier, M. (2012). The gestural approach to syllable structure: Universal, language and cluster-specific aspects. In: S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Eds.), *Speech planning and dynamics* (p. 63096). New York, Oxford, Wien: Peter Lang (Frankfurt am Main, Berlin, Bern, Bruxelles).
Saltzman, E. (1995). Intergestural timing in speech production: Data and modeling. In *Proceedings of the XIII international congress of phonetic sciences* (Vol. 2, pp. 84–91).
Saltzman, E., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*, 333–382.
Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, *3*, 371–405.
Son, M., Kim, S., & Cho, T. (2012). Supralaryngeal articulatory signatures of three-way contrastive labial stops in Korean. *Journal of Phonetics*, *40*, 92–108.
Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: Articulatory results. *Journal of the Acoustical Society of America*, *113*, 2834–2849.
Tuller, B., Shao, S., & Kelso, J. A. S. (1990). An evaluation of an alternating magnetic field device for monitoring tongue movements. *Journal of the Acoustical Society of America*, *88*, 674–679.