STABILITY OF CV INTERGESTURAL TIMING AND COORDINATION AS A FUNCTION OF PROSODIC BOUNDARY AND SYLLABLE STRUCTURE IN KOREAN

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

[a]Hanyang University, Korea; [b]Hongik University, Korea
yoon.yeomin@gmail.com; sahyang@hongik.ac.kr; tcho@hanyang.ac.kr

ABSTRACT

This study examined effects of Prosodic Boundary and Syllable Structure on the CV intergestural timing in Korean, using an EMA. The intergestural timing was more stable in the tautosyllabic (#CV) than in the heterosyllabic (C#V) condition, in line with the lexically-specified intergestural timing hypothesis, and that it was found to be more stable across an IP than a Word boundary. Nevertheless, some aspects of CV coordination remained invariant even when the consonant duration changed. Some evidence was also found, showing both dependency and independency of consonant and vowel gestures. Some of the data were interpretable in terms of gestural coupling relation.

Keywords: intergestural timing, prosodic boundary, syllable structure, Korean, CV coordination

1. INTRODUCTION

An important assumption in a mass-spring gestural model of speech production [2] is that linguistic information such as phonological contrast and lexical distinction may be expressed by differential intergestural timings—how gestures are temporally coordinated with each other. An important question is then to what extent the intergestural timing permits variability that comes from various intrinsic and extrinsic factors while maintaining the underlying phonological and lexical information. The present study explores this issue by examining how CV intergestural timing is conditioned by two interrelated factors: Syllable Structure (tauto- vs. heterosyllabic) and Prosody Boundary (IP vs. Word).

The first question is how the stability of CV intergestural timing is modified as a function of whether CV belongs to the same syllable or to different syllables across a lexical boundary. In Articulatory Phonology (AP), 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 phased than gestures that do not. While this is an important theoretical assumption of AP, only a handful of studies have tested it systematically [8, 15], and more empirical data are required to understand what it means from the kinematic point of view for gestures to be lexically specified. We therefore ask this question again, but this time we include the prosodic boundary factor, which may also influence the stability of intergestural timing through variable intergestural bonding strength [3, 8] or through gestural coupling [15]. To test this, we manipulated the syllable structure factor in such a way that the lexical effect can be concurrently examined—i.e., the tautosyllabic CV gestures occur in the same lexical item and the heterosyllabic CV gestures belong to different lexical items.

Another important question is to what extent the gestural CV coordination remains invariant and to what extent it is subject to change as the syllabic structure and the prosodic boundary vary. Previous studies have discussed cross-linguistically applicable invariant nature of CV coordination: It is not affected by the durational variation of the consonant due to an intrinsic factor such as voicing and mora [13, 14]. The present study builds on this finding by examining the durational effect of an extrinsic factor, Prosodic Boundary, to see how the boundary-induced durational change of consonants influences CV coordination. Gestural coordination is likely conditioned by a dynamic interplay between the syllable structure and the prosodic structure [6, 7]. It is therefore important to address the influence of Prosodic Boundary on CV coordination dubbed with a question as to how the boundary factor interacts with the syllable structure.

Finally, in connection with the CV coordination question, we also test CV dependency. If C and V gestures are independent from each other, operating on separate articulatory tiers [2, 9], one might expect that the kinematic realization of the vocalic gesture would not be affected by whether the consonant was in the coda position (heterosyllabic with the vowel) or in the onset position (tautosyllabic with the vowel). On the other hands, if they mutually influence each other [16], the vocalic gesture is expected to change as a function of the syllable structure. Again, the dependency issue will be explored by examining how the syllable structure interacts with the prosodic boundary factor.

2. METHOD

Electromagnetic Articulography (EMA, Carstens Articulograph, AG200) was used to obtain articulatory data. Four Seoul Korean speakers participated. For data acquisition, 8 pellets were used: two on the lower and the upper lips at the vermilion borders; three on

the tongue blade, the body and the dorsum; one at the lower gumline of the mandibular incisor; and two as reference points attached on the upper gumline and the nose bridge, respectively, which were used to correct for the head movement inside the helmet.

Table 1: Test speech materials.

Target		Target-bearing sentences
'#'=IP	#CV	[jʌŋlani ʌmʌn**i**]#[**ma**tʃʰim tʃal osjʌs*ʌjo]
		'Young-Ran's mother,#you came just in time.'
	V#C	[jʌŋlani ʌmʌn**im**]#[**a**tʃʰim tʃal mʌkʌs*ʌjo]
		'Young-Ran's mother,#I had such a good breakfast.'
'#'=Wd	#CV	[jʌŋlani ʌmʌn**i#ma**ɲil] [tʃalk*atʃitʌla]
		'Young-Ran's mother's garlic peels easily.'
	V#C	[jʌŋlani ʌmʌn**im#a**til] [tʃaŋkakas*tʌla]
		'The son of Young-Ran's mother has married.'

The test sequence was always /ima/ in which two factors were manipulated: Prosodic Boundary (IP vs. Word) and Syllable Structure (tautosyllabic #CV vs. heterosyllabic C#V). As given in Table 1, the CV sequence 'ma' in the tautosyllabic (#CV) condition occurred in the same lexical item /#**ma**tʃʰim/ ('in time') at the IP boundary and /#**ma**ɲil/ ('garlic') at the Word boundary, whereas the same sequence was separated by a lexical boundary in the heterosyllabic (V#C) condition.

Target-bearing sentences in Table 1 were repeated at least 15 times for each subject in a pseudo-randomized order. All the recorded data were carefully inspected by two trained Korean ToBI transcribers to check whether the prosodic boundary was produced as intended. Some tokens were excluded if two transcribers did not reach an agreement on the boundary type, and if there was an abnormal velocity trajectory pattern. Ten tokens were then randomly chosen to make an equal number of tokens for each condition across speakers.

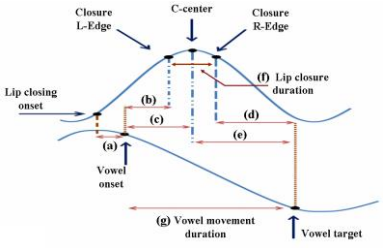
We defined articulatory landmarks using a Matlab-based kinematics analysis software, MVIEW. For the consonantal gesture, we used the kinematic data obtained by the Lip Aperture (the inter-lip Euclidean distance). As shown in Fig.1., the consonantal landmarks were *Lip Opening (Movement) Onset*, *Closure Left Edge (L-Edge)*, *C-Center*, *Closure Right Edge (R-Edge)*. For the vocalic movement, *Vowel (Movement) Onset* and *Vowel Target* were used.

To examine the stability of CV intergestural timing and coordination, we measured five temporal intervals between the gestural landmarks as illustrated in Fig. 1a-e:

(a) Vowel Onset relative to Lip Closing Onset
(b) Vowel Onset relative to Closure L-Edge
(c) Vowel Onset relative to Closure C-Center
(d) Closure R-Edge relative to Vowel Target
(e) C-Center relative to Vowel Target

Note that among the five measures, we adopted three measures (a,b,d) following Lofqvist [13]. Two intervals that involved C-center were added, as the C-center has been considered as the locus of the consonantal gesture that is phased with the vocalic gesture [1, 4]. In addition to these five intergestural timing measures, the lip closure duration (f) and the vowel movement duration (g) were included.

Figure 1: Gestural landmarks and temporal intervals between gestures.



With the five intergestural timing measures (a-e in the figure), we first obtained standard deviations with which we could assess the stability of the CV phasing relationship in a comparable way with previous studies [1, 4, 8]. Second, we analyzed intergestural intervals [13] to assess in what aspects the CV coordination remains invariant or it is subject to modification as the conditions of Syllable Structure and Prosodic Boundary change.

3. RESULT AND DISCUSSION

3.1. Stability of CV intergestural timing

Syllable structure effects. CV intergestural timing patterns were more stable when CV gestures belonged to the same syllable than to different syllables. As shown in Table 2, standard deviations were smaller in the tautosyllabic (#CV) condition than in the heterosyllabic (C#V) condition in all five measures for both IP and Word conditions. The patterns were quite consistent across four speakers, especially for three measures (a,d,e) for IP (*V-Onset to Lip Closing Onset, Closure R-Edge to V-Target, and C-Center to V-Target*), and two measures (b,c) for Wd (*V-Onset to Closure L-Edge* and *V-Onset to C-Center*). What was particularly important was that the consistent tautosyllabic stability of intergestural timing was observed in the Wd (phrase-internal) condition, although it is precisely the condition in which 'resyllabification' is expected to take place, thus being likely to neutralize #CV and C#V. Given that #CV gestures occurs in the same lexical item while C#V gestures belong to different lexical items in the present data, this syllable structure effect is consistent with the lexically specified intergestural timing account [2, 8]. This is also consistent with predictions by gestural coupling relations [15] i.e., #CV vs. C#V has in-phase (more stable) vs. anti-phase (less stable) coupling relation, which accounts for relative CV stability differences. Furthermore, since the tautosyllabic CV in the present experimental design indeed form a full syllable, the result supports the view that the syllable is a unit of articulatory organization [11], which is further in line with the notion of the syllable as the

'preprogrammed' articulatory constellation stored in the mental "syllabary" advocated by Levelt [12].

Table 2: Standard deviations as indicators of the variability of CV intergestural timing as a function of Syllable Structure.

Temporal Intervals	IP	Wd
	C#V vs. #CV	C#V vs. #CV
(a) V-Onset to Lip Closing Onset	**29.77 > 23.09 (all spkrs)**	13.7 > 9.73 (all but S3)
(b) V-Onset to Closure L-Edge	20.97 > 14.96 (all but S2)	**13.88 > 10.82 (all spkrs)**
(c) V-Onset to C-Center	17.77 > 15.61 (except S2,S3)	**14.21 > 10.84 (all spkrs)**
(d) Closure R-Edge to V-Target	**20.24 > 4.89 (all spkrs)**	11.22 > 9.06 (all but S1)
(e) C-Center to V-Target	**23.05 > 5.27 (all spkrs)**	12.32 > 8.93 (all but S1)

Table 3: Standard deviations as indicators of the variability of CV intergestural timing as a function of Prosodic Boundary.

Temporal Intervals	C#V (heterosyll)	#CV (tautosyll)
	IP vs. Wd	IP vs. Wd
(a) V-Onset to Lip Closing Onset	29.77 > 13.7 (all but S4)	**23.09 > 9.73 (all spkrs)**
(b) V-Onset to Closure L-Edge	20.97 > 13.88 (all but S4)	14.96 > 10.82 (all but S4)
(c) V-Onset to C-Center	17.77 > 14.21 (all but S4)	15.61 > 10.84 (all but S4)
(d) Closure R-Edge to V-Target	**20.24 > 11.22 (all spkrs)**	**4.89 < 9.06 (all spkrs)**
(e) C-Center to V-Target	**23.05 > 12.32 (all spkrs)**	**5.27 < 8.93 (all spkrs)**

Boundary effects. In the heterosyllabic C#V condition, all five intergestural intervals showed more stable pattern across a Wd boundary than across a IP boundary (Table 3). This is consistent with the bonding strength hypothesis [3]—i.e., the intergestural cohesiveness is inversely proportional to the size of the prosodic boundary, especially when the two gestures belong to different lexical items.

In the tautosyllabic #CV condition, the boundary effect showed a similar pattern as far as the first three intergestural intervals that involved V-Onset were concerned (a,b,c) in Table 3, showing more 'variable' phasing relationships for IP than for Word. Interestingly, however, the exact opposite pattern was observed with two other intergestural intervals that involved V-Target (d,e), showing more 'stable' intergestural timing for IP than for Wd.

Two important points can be drawn from the asymmetrical patterns. First, the consonant gesture is more stably coordinated with the right side (the target) of the vowel gesture than with its left side (the onset) in the tautosyllabic condition (the right column of Table 3), while no such difference could be observed in the heterosyllabic condition. Second, the stable phasing relationship between the consonant gesture and V-Target in the tautosyllabic condition is reinforced IP-initially. Again, this effect can be accounted for in terms of coupling relation, especially between the pi-gesture (a kind of temporal modulation gesture at a boundary) and articulatory gestures for C and V. The pi-gesture can be coupled to C and V gestures at a boundary [17]. When pi-gesture, C, and V gestures are all coupled to each other (presumably all in in-phase coupling relation), #CV coordination is expected to be more stable at a higher prosodic boundary where pi-gesture's coupling force is hypothesized to increase. This implies that the tautosyllabic intergestural timing may be further stabilized domain-initially, which can be seen as an additional characteristic of domain-initial strengthening from the gestural point of view.

3.2. CV intergestural coordination

Two way RM ANOVAs showed a significant Prosodic Boundary x Syllable Structure interaction for all five intergestural interval measures, as well as for the vowel movement duration. The lip closure duration measure also showed a trend interaction effect. These statistical results suggested that CV intergestural coordination was influenced by Prosodic Boundary and Syllable Structure factors in an interactive way, which allowed us to focus on the results related to the interaction effects (Table 4) in light of research questions of the present study.

Table 4: A summary of planned pair-wise comparisons by a series of t-tests. ('*' refers to $p<0.05$, '**', $p<0.01$.)

Temporal Measures	Boundary Effect t(3)		Syll. Effect t(3)	
	C#V	#CV	IP	Wd
(a) V-Onset to Lip Closing Onset	IP>Wd t=-4.3*	**IP=Wd** t=-1.0	C#V>#CV t=-24.7**	**C#V=#CV t=-1.4**
(b) V-Onset to Closure L-Edge	**IP=Wd t=-1.0**	**IP=Wd** t=0.3	C#V<#CV t=-5.6*	**C#V=#CV t=-0.9**
(c) V-Onset to C-Center	**IP=Wd t=-0.2**	**IP=Wd** t=0.7	C#V<#CV t=-3.5*	**C#V=#CV t=-1.1**
(d) Closure R-Edge to V-Target	IP>Wd t=-4.0*	**IP=Wd** t=1.1	C#V>#CV t=-4.9*	**C#V =#CV t=-0.2**
(e) C-Center to V-Target	IP>Wd t=-5.1*	**IP=Wd** t=0.3	C#V>#CV t=-6.1**	**C#V=#CV t=0.1**

CV coordination difference by Boundary. Let us first look at the boundary effect on the intergestural intervals in the C#V condition. Two intervals that involved V-Target, (d,e), showed that the distance between a consonantal landmark and V-Target was longer for IP than for Wd: V-Target came later at the IP boundary than at the Wd boundary for C#V. On the other hand, two intervals that involved V-Onset and consonant closure landmarks (b,c), were not affected at all by the Boundary factor, showing that the invariant CV coordination was maintained at least between a consonant closure landmark (*L-Edge* or *C-Center*) and V-Onset. So it appears that the right side of the vowel gesture becomes farther away from the consonantal gesture in C#V across an IP boundary, reinforcing the heterosyllabicity of the two gestures as kind of domain-initial strengthening, while the invariant CV coordination is maintained between the left side of the vowel gesture and the consonant. Note that the invariance was observed even though the

closure duration was modified as a function of Boundary, as given in Table 5.

The invariant CV coordination crystallizes in the tautosyllabic (#CV) condition. While the closure duration varied with the boundary condition (Table 5), all five intergestural intervals remained unchanged for IP vs. Wd (Table 4). This demonstrates that CV coordination remains invariant even when the closure duration is modified by an extrinsic, prosodic boundary factor, in a comparable way with the case of the invariant CV coordination reported for variable consonant closure duration due to intrinsic phonemic differences [13, 14].

CV coordination differene by Syllable Structure. The rightmost two columns of Table 4 showed how CV coordination changes as a function of Syllable Structure. An interesting pattern was observed in the IP condition: For intervals involving V-Onset (b-c in Table 4), they were longer for C#V than #CV, but for intervals involving V-Target (d,e), the opposite was observed—i.e., they were longer for #CV than C#V. In other words, the consonant gesture came relatively earlier, being shifted to the left side for C#V in Table 4 (b,c), but shifted to the right side for #CV in Table 4 (d,e), supporting the view that the C and V gestures operate on separate tiers [9]. Another interesting point is that V-Onset to Lip Closing Onset in Table 4(a) is shorter for #CV than for C#V at least in the IP condition. This can be accounted for by the coupling relation: C and V onsets are expected to be closer to each other in an in-phase coupling relation (#CV) than in an anti-phase coupling relation (C#V) [10].

C-V dependency. Nevertheless, the V movement duration was found to be longer for C#V than for #CV in the IP condition (Table 5), suggesting that at least at the IP boundary, the vocalic movement was not entirely independent from the consonant gesture. Taken together, the results imply that the vocalic and consonantal gestures may be independent, as reflected in the 'sliding' effect as a function of Syllable Structure, but that the vocalic gesture can be independently lengthened for C#V across an IP boundary. The independent lengthening can be interpreted as a kind of 'vocalic' domain-initial strengthening effect, given that 'V' in C#V is under a direct influence of domain-initial strengthening.

Table 5: A summary of planned pair-wise comparisons by a series of t-tests for C and V durations ('*', $p<0.05$, 'tr', $p<0.09$).

Temporal Measures	Boundary effect		Syll. effect	
	Coda	onset	IP	Wd
(f) Lip Closure Duration	**IP>Wd t=5.3***	**IP>Wd t=3.7***	**C#V>#VC t= 2.6tr**	C#V =#VC t=-1.6
(g) Vowel Duration	**IP>Wd t=2.5tr**	IP=Wd T=0.7	**C#V>#VC t=5.7***	C#V =#VC t=0.7

4. CONCLUSION

The implications of the present study can be pinpointed as follows. First, the stable intergestural timing in the tautosyllabic condition supports the hypotheses about the lexically-specified intergestural timing and the syllable as a physiological unit, which may be further interpretable by gestural coupling relation. Second, CV intergestural timing across a boundary (C#V) was more variable, suggesting that the gestural cohesiveness becomes weaker for a larger boundary. Third, some CV coordination patterns were not conditioned by the boundary-induced durational change of the consonant, demonstrating the invariant nature of CV coordination. Fourth, some evidence was found for both CV gestural dependency (i.e., V duration was affected by the syllable structure) and independency (i.e., the sliding effect of the consonant gesture on a separate tier).

5. REFERENCES

[1] Browman, C., Goldstein, L. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140-155.

[2] Browman, C., Goldstein, L. 1992. Articulatory phonology: an overview. *Phonetica* 49, 155-180

[3] Browman, C., Goldstein, L. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parle* 5, 25-34.

[4] Byrd, D. 1995. C-center revisited. *Phonetica* 52, 263-282.

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

[6] 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 Fougeron, C., Kühnert, B., d'Imperio, M., Vallée, N. (eds.), *Phonology and Phonetics: Laboratory Phonology.* De Gruyter, Berlin, 10, 31-59.

[7] Byrd, D., Krivokapic, J., Lee, S. 2006. How far, how long: On the temporal scope of phrase boundary effects. *J. Acoust. Soc. Am.* 120, 1589-1599.

[8] Cho, T. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58, 129-162.

[9] Fowler, C.A. 1983. Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology* 112, 386-412.

[10] Goldstein, L., Byrd, D., Saltzman, E. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In Arbid, A.M. (ed.), *From Action to Language: The Mirror Neuron System*. Cambridge: Cambridge University Press, 215-249.

[11] Krakow, R.A. 1999. Physiological organization of syllables: A review. *Journal of Phonetics* 27, 23-54.

[12] Levelt, W.J.M. 1989. *Speaking: From Intention to Articulation.* Cambridge: The MIT Press.

[13] Löfqvist, A. 2006. Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *J. Acoust. Soc. Am.* 120, 2872-2883.

[14] Löfqvist, A., Gracco, V.L. 1999. Interarticulator programming in VCV sequences: Lip and tongue movements. *J. Acoust. Soc. Am.* 105, 1864-1876.

[15] Nam, H. 2007. Syllable-Level intergestural timing model: Split-Gesture dynamics focusing on positional asymmatry and moraic structure. In Cole, J., Hualde, I.J. (eds.), *Phonology and Phonetics: Laboratory Phonology*. New York: Mouton de Gruyter, 9, 483-506.

[16] Öhman, S.E.G. 1967. Numerical model of coarticulation. *J. Acoust. Soc. Am.* 41, 310-320.

[17] Saltzman, E., Nam, H., Krivokapic, J., Goldstein, L. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proc. Speech Prosody* Campinas, 175-184.