Intergestural CV timing of homophonous words with different morphological structures: A preliminary report on liquid /l/ in Korean

Jiyoung Lee[1], Sahyang Kim[2], and Taehong Cho[1]
[1]*Hanyang Institute for Phonetics & Cognitive Sciences of Language, Hanyang University (Korea),*
[2]*Hongik University (Korea)*
Ljy1004@hanyang.ac.kr , sahyang@hongik.ac.kr, tcho@hanyang.ac.kr

The Korean liquid phoneme /l/ has several allophones, which are highly dependent on its position in syllable (e.g., a lateral in coda and a flap in onset or intervocalic position) [1]. This EMA study investigates whether and how the same allophone of /l/ may be distinguished in the articulatory dimension when it has different underlying syllable compositions. To examine this question, homophonous sequences with different underlying structures were compared. Two pairs of C_1V_1**$\underline{C_2}$**-C_1V_1.**$\underline{C_2}$**V_2 words (e.g., /pal/ 'foot' - /pali/ 'bowl' where /l/ is underlyingly in the coda or in the intervocalic onset position) were used, and they were followed by the grammatical particle /(i)lago/. In this way, C_2, /l/, in the words with different syllable structures would both be produced as a flap at the surface phonetic level (e.g., a derived flap condition as in /$\underline{\text{pal+i}}$lago/ and an underlying flap condition as in /$\underline{\text{pali}}$+lago/).

A couple of competing predictions can be made on the production of the two types of flaps. One possibility is that there may be no articulatory differences between the two homophonous sequences. This is based on previous studies suggesting that the spelled-out segments go through the same process and create the syllabified form, resulting in little chance to reflect its internal syllable structure in production [2]. Alternatively, some differences may be observed between the derived versus the underlying flap. In Articulatory Phonology (AP) [3, 4], gestural coordination is specified in the lexicon, and understood by phase relations within a syllable. AP assumes that the CV sequence shows an in-phase relationship, and the two gestures start almost synchronously. Based on the assumption, the underlying flap in /pali/ should show an in-phase relationship while the resyllabified flap in /pal+i/ may not show the same pattern as its relationship, which is not specified in the lexicon. Some differences in gestural coordination depending on underlying compositions have indeed been found in previous studies [5, 6]. This study further tests whether and how prosodic prominence modulates the potential underlying structural differences. It is well-known that the presence of prominence contributes to maximizing lexical contrasts and enhancing the gestural bonding strength [7, 8, 9, 10]. It is therefore hypothesized that, if any, the structural differences would be maximized under prominence.

Articulatory data were collected using EMA (AG501, Carstens Electronics) from twelve native Seoul Korean speakers but only a subset of the data was analyzed here (4F, 3M). As shown in Table 1, two prosodic factors were manipulated: Boundary (IP-initial or Wd-initial) and Focus (Focused or Unfocused). Each speaker produced 240 test sentences (4-target * 2-boundary * 2-focus * 15-repetition). By excluding 45 tokens with unintended prosodic renditions, 1635 tokens were collected for analysis. Kinematic data from tongue tip (TT) for the consonantal /l/ gesture and tongue body (TB) for the vocalic gestures were analyzed by using MVIEW [11]: C_2 duration (from the onset to the target of C_2 /l/), C_2V_2 duration (from the onset of C_2(or V_2[i]) to the target of V_2), Intergestural timing (between the onsets of C_2 and V_2[ii]). A series of linear mixed-effects models were fitted separately for each measurement.

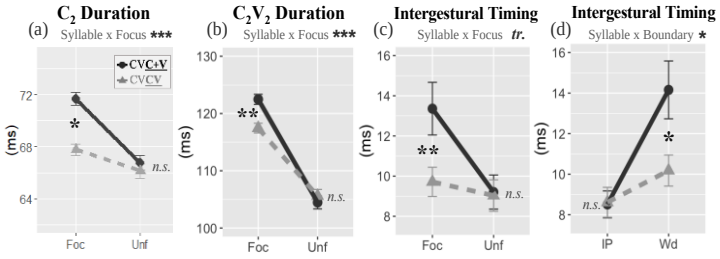
Results showed that there was a temporal difference in the articulatory duration of TT movement for /l/ as a function of underlying structure, but only in the focused condition. While /l/ was realized as a flap on the surface in both conditions, the resyllabified flap ($C_1V_1\underline{C_2}$+V_2), a lateral underlyingly, was longer than the canonical flap ($C_1V_1.\underline{C_2}V_2$) under focus (Fig. 1a, p=.028). This is in line with the lexical contrast maximization under prominence reported in previous studies [8, 10]. The results, therefore, seem to indicate that when there is a need to deliver an informational locus, speakers put a deliberate effort into making a distinction between the two homophonic sequences by referring to the

underlying syllable structures. Another important finding was that C_2V_2 duration was shorter under focus in the monomorphemic ($C_1V_1.\underline{C_2V_2}$) than in the heteromorphemic condition ($C_1V_1\underline{C_2+V_2}$) (Fig. 1b, p=.0029), showing underlying syllable differences. In addition, the absolute distance between the C_2 and V_2 gestures (i.e., intergestural timing between C_2 and V_2 gestures) was again shorter for $C_1V_1.\underline{C_2V_2}$ than for $C_1V_1\underline{C_2+V_2}$ in the focused condition (Fig. 1c, p=.009). According to Articulatory Phonology [3, 4], the gestural coordination and corresponding phase relationship is specified in the lexicon, and C and V gestures are assumed to start synchronously as they are in an in-phase relationship. Thus, the smaller temporal interval between C and V gestures can be understood as a tighter in-phase relationship with stronger gestural cohesiveness. The differences were further augmented under prosodic prominence, suggesting that speakers make efforts to differentiate the underlying structural difference. What is interesting is the difference in intergesutral timing between the two homophonous sequences was found in the Wd condition. A stronger prosodic boundary is generally known to induce a stronger bonding relationship between C and V [7, 9], resulting in shorter temporal interval in the IP-initial condition (Fig 1d). On the contrary, in the phrase-medial position, loosened gestural bonding strength may display the underlying structural differences, resulting in a longer temporal interval between C and V gestures especially for $C_1V_1\underline{C_2+V_2}$ than in $C_1V_1.\underline{C_2V_2}$ (Fig 1d, p=.02). In conclusion, the present study suggests that speakers fine-tune the articulatory realization of gestures and their coordination to encode and maintain the underlying structural difference, further modulating them by referring to higher-order prosodic structure.

Table 1. Examples of test sentences. Targets are underlined, and contrasted words are marked in bold. '#' and '+' refer to phrase boundary and morphological boundary, respectively.

Conditions		Test sentences
#=IP	Foc	/ʧikimjʌki, # **pali**+lago s*ʌnni, # **pal+**ilago s*ʌnni/ Right here, did you write a **bowl** or a **foot**?
	Unf	/ʧikimjʌki, # pal+ilago **nika** s*ʌnni, # pal+ilago **ʧjeka** s* ʌnni/ Right here, did **you** write a foot or did **that person** write a foot?
#=Wd	Foc	/ʧikimjʌki, wuli#**pali**+lago s*ʌnni, wuli#**pal+**ilago s*ʌnni/ Right here, did you write our **bowl** or our **foot**?
	Unf	/ʧikimjʌki, wuli#pal+ilago **nika** s*ʌnni, wuli#pal+ilago **ʧjeka** s* ʌnni/ Right here, did **you** write our foot or did **that person** write our foot?

Figure 1. Syllable x Focus interaction on C_2 duration (a), C_2V_2 duration (b), and intergestural timing (c). Syllable x Boundary interaction on intergestural timing (d). Error bars represent standard errors ('*' refers to p<.05, '**' to p<.01, and '***' to p<.001).



References

[1] Lee, S. & Kang, S. (2003). Acoustic and phonological properties of Korean liquids. Journal of Language Sciences, 10(2), 77-92.

[2] Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. Behavioral and brain sciences, 22(1), 1-38.

[3] Browman, C.P., & Goldstein, L. (1988). Some notes on syllable structure in Articulatory Phonology. Phonetica, 45, 140–155.

[4] Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180.

[5] Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. Phonetica, 58(3), 129-162.

[6] Lee, J., Kim, S., & Cho, T. (2019). Effects of morphological structure on intergestural timing in different prosodic-structural contexts in Korean. In Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia.

[7] Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. Journal of Phonetics, 29(2), 155-190.

[8] Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. Journal of Phonetics 39.

[9] Cho, T., Yoon, Y., & Kim, S. (2014). Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures. Journal of Phonetics, 44, 96-109.

[10] de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. The journal of the acoustical society of America, 97(1), 491-504.

[11] Tiede, M. (2010). MVIEW: Multi-channel visualization application for displaying dynamic sensor movements.

[i] This is the case when the V_2 gesture starts earlier than the C_2 gesture.

[ii] In order to measure the synchronicity of the two gestures, the actual degree of proximity between the onsets of C_2 and V_2 was measured as an absolute value.