ARTICULATORY MANIFESTATION OF FOCUS-INDUCED PROMINENCE IN SEOUL KOREAN

Jiyoung Jang & Argyro Katsika

University of California, Santa Barbara
jiyoung@ucsb.edu, argyro@ucsb.edu

ABSTRACT

Speech units under phrase-level prominence are known to be produced with longer, larger, and faster constriction gestures. Evidence, however, comes mainly from head-prominence languages. The present electromagnetic articulography study investigates the effect of focus-induced prominence on articulatory manifestation in Seoul Korean, an edge-prominence language that marks phrase-level prominence via Accentual Phrases (APs) and also lacks word-level prosody. Special interest is given in examining the scope of the effect as well as its interaction with boundary marking at the level of the Intonational Phrase (IP). Results from seven native Seoul Korean speakers suggest that constriction gestures are longer, larger, and faster under focus. The scope of the effect extends away from the initial syllable of the focused AP reaching up to its third syllable. An interaction between focus and final IP boundary is detected, indicating an intricate relationship between information structure and prosodic structure, discussed within the framework of Articulatory Phonology.

Keywords: focus-induced prominence, articulation, prosody, Seoul Korean, Articulatory Phonology

1. INTRODUCTION

Prosodic structure serves two important linguistic functions in speech: phrasing, which groups linguistic units into larger cognitive constituents, and prominence, which marks constituents that are important for either rhythmic (e.g., via stress) or conceptual reasons (e.g., via pitch accent). Work on the phonetic realization of these prosodic landmarks has been provided accumulated evidence suggesting that segments are produced with longer, larger, and faster constriction gestures when accented and with longer, less overlapped and slower constriction gestures when phrase-final (e.g., [1, 2]). However, these results come mainly from head-prominence languages, such as English, in which phrase-level prominence is usually marked by placing a pitch accent on the stressed syllable of the prominent word (cf. [3]).

Here, we turn to Seoul Korean with the goal to examine the articulatory correlates of focus-induced prominence in an edge-prominence language. Seoul Korean does not employ word prosody (i.e., lexical stress, lexical tone or lexical pitch accent), and phrase-level prominence is marked by the means of Accentual Phrases (APs). AP serves as the basic intonational unit and is marked by a particular pitch contour [4, 5]. AP's proposed underlying tonal pattern is THLH, where the realization of the initial tone (T) tends to depend on the laryngeal configuration of the AP-initial segment [4 - 7]. Focus-induced prominence in Korean employs this AP level, since the focused word consistently starts an AP (or a higher phrase), and any following AP boundaries up to the end of the Intonational Phrase (IP) often undergo elimination, or possibly attenuation, referred to as dephrasing [6, 8].

Limited work on the correlates of prominence in Korean reports longer, larger, and slightly faster vocalic movements under focus in Korean [9] (see [10] for a review). However, the scope of the effect as well as its interaction with marking higher-level IP boundaries is unclear. The latter is of interest, because in Korean both phrasing and prominence are marked via prosodic boundaries. In particular, we examine the interaction between focus marking (left-edge AP boundary) and IP marking (right-edge IP boundary), with a goal to broaden our understanding on the interplay between information structure and prosodic structure. Another interesting research question involves the phenomenon of dephrasing that arises from the effect of focus-induced prominence; the current study also examines whether the patterns of dephrasing differ with distance of focus from the right edge of the IP.

We use Electromagnetic Articulography (EMA) to assess the effect of contrastive focus on the duration, displacement, and velocity of the consonant gestures of the focused words. We expect gestures under focus to be longer, larger, and faster as compared to that of unfocused ones, in line with previous findings [9]. In addition to these articulatory manifestations, we investigate the relationship between kinematic parameters. Specifically, we test whether duration increases as stiffness decreases, and whether displacement increases proportionally to velocity, as it has been shown to be the case in previous research primarily on head-prominence languages [11, 12].

Based on findings of the extended scope of prominence effect beyond stressed syllable (head) in head-prominence languages [13], we expect the scope of the focus-induced prominence to span more than the initial syllable of the AP. Nonetheless, to which syllable the effect extends to as well as whether it would interact with the effect spanning from the IP boundary marking from the right-edge needs to be tested.

2. METHODS

2.1. Participants and experimental procedure

Seven native Seoul Korean speakers (5F, 2M; Mean age = 24.5, Age range = 21-28) participated in the study. The speakers were naïve as to the purpose of the study and had no reported speech, hearing, or vision problems. They received financial compensation for their participation.

Prior to the experiment, the participants went through a 15-minute training session in order to become acquainted with the speech materials and their presentation. Kinematic data were collected using an AG501 3D electromagnetic articulograph (Carstens Medizinelektronik). Ten receiver coils were attached to the tongue dorsum (two sensors), tongue tip, upper and lower incisors, upper and lower lips, left and right ears, and nose. Audio recordings were performed simultaneously to the kinematic recordings by the means of a Sennheiser shotgun microphone positioned approximately one foot away from the participant's mouth and set at a sampling rate of 16 kHz. Stimuli were presented on a computer screen placed roughly three feet away from the participant. The experimental session began with a set of simple tasks (e.g., reading lists of short words), the goal of which was to familiarize the participants with the experimental procedure and apparatus. Next, participants were asked to read sets of prompt-target sentences (see Section 2.2). For each set, the prompt sentence was shown in green font and appeared first. The target sentence, shown in blue font, appeared one second after the prompt sentence, and was read aloud. Both prompt and target sentences were presented in regular font, i.e., non-bolded and non-underlined.

2.2. Stimuli

The effect of focus was examined across the test interval /nɛ.maŋ.mi.nam/, a compound word that means *'a handsome guy from Nemang'*. The participants were asked to imagine a situation where they were preparing a play and the compound word was presented as a role in the play. The target word was framed to receive contrastive focus in a set of test stimuli, as shown in Table 1a. Prompt sentences were designed to help appropriate focus placement. For example, for the target sentence (a) in Table 1, the prompt sentence was *'It's not the handsome guy from Nowon (that Uncle Minam is playing)'*. Focused test words in the stimuli were compared to two sets of control stimuli that included the test word in unfocused condition; in one set, focus was placed on the second word of the sentence (Table 1b), and in the other, focus was on the first word of the sentence (Table 1c). We refer to these conditions as Focus-proximate and Focus-distant, respectively. Having these two different unfocused conditions allows examining the degree of dephrasing.

To assess the interaction between focus marking and IP-boundary, test stimuli were further manipulated so that test words either IP-final or IP-medial (compare stimuli (a) and (d) in Table 1).

The combination of Focus Type (Focused, Focus-proximate, Focus-distant) and Within-IP Position (IP-final, IP-medial) yielded six conditions in total, which occurred in an experimental block along with stimuli for other experiments. Each block was differently randomized and repeated eight times. Note that for one speaker, five repetitions were collected due to interruption of the experimental session for technical reasons. In total, 318 tokens were included in the analyses reported here. The acquired data were checked for their prosodic rendition, i.e., focus placement and prosodic boundaries, using K-ToBI [14].

Focus	***Target sentence (IP-final)***
(a) Focused	[pimilpu minami gomopuga **nɛmaŋminam**]? [sʌntækhangʌja]? 'Uncle Minam of the secrecy club is **the handsome guy from Nemang**? Is it decided?'
(b) Focus-proximate	[pimilpu **minami** gomopuga nɛmaŋminam]? [sʌntækhangʌja]? 'Uncle **Minam** of the secrecy club is the handsome guy from Nemang? Is it decided?'
(c) Focus-distant	[**pimilpu** minami gomopuga nɛmaŋminam]? sʌntækhangʌja]? 'Uncle Minam of **the secrecy club** is the handsome guy from Nemang? Is it decided?'
Focus	***Target sentence (IP-medial)***
(d) Focused	[pimilpu minami gomopuga **nɛmaŋ**minam sʌntækhangʌja]? 'Uncle Minam of the secrecy club chose **the handsome guy from Nemang**?'

Table 1: Example stimuli presented by Focus Type (Focused, Focus-proximate, Focus-distant). The IP-final and IP-medial counterparts are given for the Focused condition.

2.3. Data analysis

Consonant (C) gestures in the interval /nɛ.maŋ.mi.nam/ were analyzed, except the coda of the second syllable. Coda /ŋ/ was excluded from the analysis because of its degree of blending with the neighboring vowels. We will refer to the measured consonants as C1, C2, C3, C4, and C5. All of these were onsets, except for C5 which is the coda of the fourth syllable. Kinematic labeling was performed using semi-automatic custom software in Matlab (Mark Tiede, Haskins Laboratories).

For labelling, tongue tip vertical displacement trajectory was used for /n/ and lip aperture trajectory was used for /m/. The labelling procedure detected the following kinematic timepoints in each C gesture on the basis of velocity criteria: onset, time of peak velocity, target, constriction maximum, release, and offset (Figure 1). The following kinematic measures were calculated based on these timepoints: formation duration (interval between onset and release), displacement (spatial difference between max and onset), and formation's peak velocity (velocity value at point (c) in Figure 1) of each test C gesture.

The retrieved data were analyzed by linear mixed effects analysis using lme4 [15] package in R [16]. To test the effect of focus type and its interaction with Within-IP Position, each C gesture was examined for the dependent variables of formation duration, displacement, and peak velocity. Fixed effects of Focus type (Focused, Focus-proximate, Focus-distant) and Within-IP Position (IP-final, IP-medial) were included. Random effect of speaker was added. To examine the relationship between kinematic parameters [11, 12], two models were built. One model was with formation duration as dependent variable, and stiffness (as calculated by peak velocity over displacement [17]) and Consonant (C1 to C5) as fixed structure. Consonant was added to see whether there was difference in the relationship between the parameters as a function of consonant position in the word or phrase. The other model had displacement as dependent variable, and peak velocity and Consonant as fixed structure. Random effect of speaker was added in the models. In case of significant effects, pair-wise comparisons were assessed by the *emmeans* package [18] with Holm adjustment.

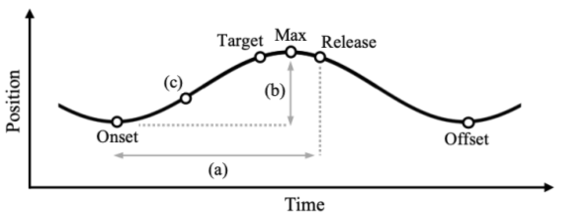

Figure 1: Kinematic measurements of constriction gesture: (a) formation duration (in ms), (b) displacement (in mm), and (c) peak velocity (in cm/sec).

3. RESULTS

3.1. Effect of focus type and its interaction with phrase-final boundary

Table 2 lists the significant main and interaction effects of Focus Type and Within-IP Position, and Figure 2 visualizes the results. Main effect of Focus Type was found in the C gestures of the three first syllables of the test word (C1, C2, C3; Table 2, lines 1-9). In C1, C gestures were longer and larger in the Focused condition than in either of the Unfocused conditions ($p<0.001$ for all comparisons). Gestures were faster when focused as opposed to when focus fell on the proximate word ($p<0.05$), but not when it fell on the distant word ($p>0.09$). No significant difference was found between the two Unfocused conditions in any of the measurements.

Syll	*DV*	*Focus*	*Boundary*
C1	(1) duration	$F(2)$=98.8***	*n.s.*
	(2) displacement	$F(2)$=23.1***	*n.s.*
	(3) peak velocity	$F(2)$=6.6*	*n.s.*
C2	(4) duration	$F(2)$=51.0***	*n.s.*
	(5) displacement	$F(2)$=110.3***	*n.s.*
	(6) peak velocity	$F(2)$=62.7***	*n.s.*
C3	(7) duration	n.s.	*n.s.*
	(8) displacement	$F(2)$=164.5***	*n.s.*
	(9) peak velocity	$F(2)$=114.7***	*n.s.*
C4	(10) duration	*n.s.*	$F(1)$=19.5***
	(11) displacement	*n.s.*	$F(1)$=9.6**
	(12) peak velocity	*n.s.*	$F(1)$=9.3**
C5	(13) duration	*n.s.*	$F(1)$=331.6***
	(14) displacement	*n.s.*	$F(1)$=375.4***
	(15) peak velocity	*n.s.*	$F(1)$=66.6***

Table 2: Summary of main effects ('***' = $p<0.001$, '**' = $p<0.01$, '*' = $p<0.05$, 'n.s.' = $p>0.09$).

In C2, C gestures were longer, larger and faster when focused as opposed to unfocused ($p<0.001$ for all comparisons).

In C3, formation duration did not show systematic difference as a function of Focus Type. However, both displacement and peak velocity did, with C gestures being larger and faster when focused as compared to unfocused ($p<0.001$ for all comparisons). There was also a significant interaction between Focus Type and Within-IP Position on the peak velocity of C3 ($F(2)$=9.9, p<0.01). Pairwise comparisons detected a significant difference between the two unfocused conditions but only in the IP-medial position. Specifically, IP-medially, C3 gestures were faster when focus fell on the distant word ($p<0.05$) as opposed to the proximate word ($p>0.09$).

No further main or interaction effect of Focus Type was found beyond C3. In C4 and C5, there were main effects of Boundary Type in all measurements such that gestures were longer, larger, and faster in IP-final than in IP-medial positions (Table 2, lines 10-15).

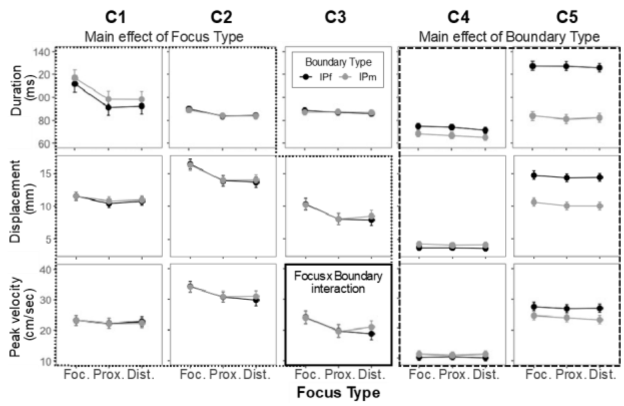

Figure 2: Formation duration (in ms), displacement (in mm), and peak velocity (in cm/sec) of C gestures (C1 to C5) as a function of Focus Type and Within-IP Position. Dotted and dashed boxes indicate main effects of Focus Type and Within-IP position, respectively, and solid boxes indicates interaction between the two factors.

In sum, consonant gestures are longer, larger, and faster under focus. The effect of focus-induced prominence extends rightward to the onset of the third syllable (C1, C2, C3). The effect of IP-boundary marking extends leftward to the onset of the fourth syllable (C4, C5). Interaction between focus and IP-final boundary is detected in the middle, and specifically on peak velocity of C3.

3.2. Relationship between kinematic parameters

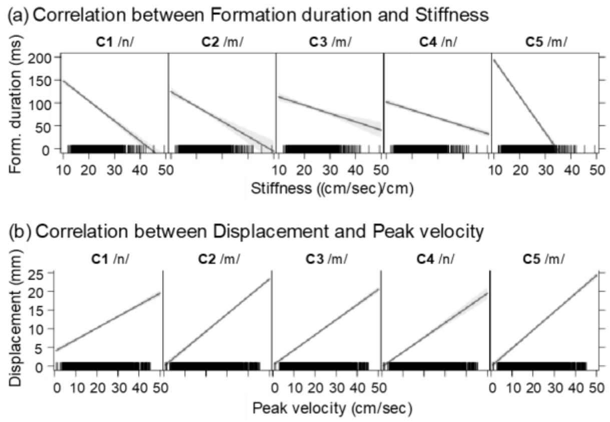

Figure 3: Relationship between (a) formation duration (in ms) and stiffness (in (cm/sec)/cm); (b) displacement (in mm) and peak velocity (in cm/sec) by Consonant.

Figure 3a demonstrates the relationship between formation duration and stiffness by Consonant. Expected covariance between formation duration and stiffness was observed ($F(1)=1100.5$, $p<0.001$) such that duration decreased as stiffness increased (cf. [11, 12]). Consonant interacted with stiffness ($F(4)=452.0$, $p<0.001$). As shown in Figure 3a, the interaction stemmed in part from the fact that the slopes were steeper when the gestures were at the edges of the phrase. Figure 3b shows the relationship between displacement and peak velocity. Expected covariance between the parameters were observed ($F(1)=3568.3$, $p<0.001$), such that displacement increased as peak velocity increased [11, 12]. Consonant interacted with peak velocity ($F(4)=98.1$, $p<0.001$), with slopes being shallower for /n/ than /m/.

4. DISCUSSION

The goal of the present study was to investigate the effect of focus-induced prominence on articulatory modulation in Seoul Korean. We found that constriction gestures under focus are longer, larger, and faster than unfocused counterparts, in line with previous findings on Korean [9]. This kinematic profile of focus is similar to that of head-prominence languages [1], despite the typological difference. One of the main findings of the present study is that the effect of focus-induced prominence goes beyond the edge of the AP initial syllable, extending rightward to the onset of the third syllable of the focused AP. Moreover, an interaction between focus position and Within-IP position was detected, indicating an intricate relationship between phrase-level prominence marking and IP boundary marking. In Articulatory Phonology [19], such prosodic modulations are instantiated by modulation gestures, or μ-gestures [20], that are either spatial or temporal. One specific type is the π-gesture, which is essentially a temporal μ-gesture marking boundaries. μ-gestures have been proposed to alter the spatiotemporal profile of the constriction gestures that overlap with them, with the scope of the effect depending on their coordination [20]. Based on our results, Seoul Korean combines a temporal μ-gesture and a spatial one to instantiate phrase-level prominence, coordinated with the beginning of the focused AP. The temporal μ-gesture may have a shorter activation interval than the spatial one, as reflected in the onset of the third syllable which showed larger and faster movement under focus, but not necessarily longer in duration. The IP-marking π-gesture seems to extend at least two syllables from the right edge of the phrase, overlapping with the μ-gesture in the third syllable. Note that constriction gestures were longer, larger, and faster when preceded by an IP boundary (see also [21]). The faster movement under the activation of π-gesture is interesting, which could possibly be arising from the interaction between π-gesture and μ-gesture.

Another question of the current study was to test whether the distance of focus would exhibit different patterns of dephrasing. Indeed, the dimension of velocity was affected by focus placement: gestures were the slowest when focus fell on the proximate word. This may suggest that the degree of dephrasing might be greater near the focused item and decrease with distance from it.

Finally, we examined the relationship between the kinematic parameters, based on previously reported correlations [11, 12]. Our results showed expected covariance between parameters—duration increases with decreased stiffness and displacement increases with peak velocity, in line with findings of head-prominence languages. However, an interesting pattern arose as a function of consonant position. The relationship between duration and stiffness showed steeper slope, i.e., duration increases more given the same decrease of stiffness, at the edges of the boundaries compared to the medial positions of the phrase. This result may be attributed as a trait of an edge-prominence language, as these positions are expected to be the main locus of focus marking (left-edge) and IP boundary marking (right-edge).

5. ACKNOWLEDGEMENTS

We would like to thank the Korean speakers who participated in the study. This work was supported by NSF Grant #1551428 and funds from the UCSB Academic Senate.

6. REFERENCES

[1] Cho, T. 2006. Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in Laboratory Phonology VIII: Varieties of Phonological Competence (Phonology and Phonetics)* (pp. 519–548). Berlin, Germany: Mouton de Gruyter.

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

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

[4] Jun, S. A. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15(2), 189-226.

[5] Jun, S. A. 2005. Korean intonational phonology and prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing*, 1, 201.

[6] Jun, S. A. 1993. *The Phonetics and Phonology of Korean*. PhD dissertation, The Ohio State University.

[7] Jun, S. A. 2006. Intonational Phonology of Seoul Korean Revisited. *UCLA Working Papers in Phonetics*, 104, 14-25.

[8] Jeon, H. S., & Nolan, F. 2017. Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology*, 8(1), 1-30.

[9] Shin, S., S. Kim, and T. Cho. 2015. What is special about prosodic strengthening in Korean: evidence in lingual movement in V#V and V#CV. *Proc. 17th ICPhS*, Glasgow, U.K.

[10] Cho, T. 2022. The phonetics-prosody interface and prosodic strengthening in Korean. *The Cambridge Handbook of Korean Linguistics Cambridge Handbooks in Language and Linguistics* (pp. 248-293). Cambridge University Press Cambridge.

[11] Edwards, J., Beckman, M. E., & Fletcher, J. 1991. The articulatory kinematics of final lengthening. *J. Acoust. Soc. Am.*, 89, 369–382.

[12] Byrd, D., & Saltzman, E. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31(2), 149–180.

[13] Katsika, A., & Tsai, K. 2021. The supralaryngeal articulation of stress and accent in Greek. *Journal of Phonetics,* 88, 101085.

[14] Jun, S. -A. 2000. K-ToBI (Korean ToBI) labelling conventions. *Speech Sciences*, 7(1), 143-170.

[15] Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

[16] R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[17] Ostry, D. J., & Munhall, K. G. 1985. Control of rate and duration of speech movements. *J. Acoust. Soc. Am.*, *77*(2), 640-648.

[18] Lenth, R. V. 2020. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.3. https://CRAN.R-project.org/package=emmeans.

[19] Browman, C. P., & Goldstein, L. M. 1992. Articulatory phonology: an overview. *Phonetica*, 49, 155–180.

[20] Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proc. 4th Speech Prosody 2008*, Campinas, Brazil, 175-184.

[21] Kim, J. J., Baek, Y., Cho, T., & Kim, S. 2019. Preboundary lengthening and boundary-related spatial expansion in Korean. *Proc. 19th ICPhS*, Melbourne, Australia, 1019-1023.