SEPTEMBER 25 2025

Kinematic encoding of focus and edge-prominence in Seoul Korean

Jiyoung Jang; Argyro Katsika

J. Acoust. Soc. Am. 158, 2523–2536 (2025)
https://doi.org/10.1121/10.0039347

JASA ARTICLE

Kinematic encoding of focus and edge-prominence in Seoul Korean

Jiyoung Jang[1,a)] and Argyro Katsika[2]
[1]*Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul 04763, South Korea*
[2]*Department of Linguistics, University of California, Santa Barbara, Santa Barbara, California 93106, USA*

ABSTRACT:
Articulatory gestures under phrasal prominence undergo strengthening, becoming longer, larger, and faster. Limited research, mainly on head-prominence languages, shows that prominence-induced strengthening interacts with focus structure, increasing gradually across focus types. However, it is unclear how focus structure is encoded in edge-prominence systems. Here, we examine Seoul Korean, an edge-prominence language, in which the focused word is known to start an accentual phrase (AP) and exhibits prominence-induced strengthening, while the post-focal items are dephrased. Analyses of kinematic duration, displacement, and velocity, examine degree of strengthening on focused AP-initial gestures and/or dephrasing on initial gestures in the first post-focal word. Results show that focused AP-initial strengthening reflects focus structure, although kinematic dimensions differ in the number of focus types they distinguish. Yet, the order of encoded types remains consistent and similar to that found in head-prominence languages. Post-focally, there is durational evidence of dephrasing only after contrastive focus and its reach is constrained by the number of intervening syllables. Instead, the other focus types exert strengthening on the onset of the post-focal word, suggesting focus-induced spillover effects that cross-word boundaries. These findings support the view that prominence is organized as a hierarchical structure, with its levels reflecting different focus types. https://doi.org/10.1121/10.0039347

(Received 1 March 2025; revised 26 August 2025; accepted 28 August 2025; published online 25 September 2025)
[Editor: Jianjing Kuang] Pages: 2523–2536

I. INTRODUCTION

Prosodic structure plays a key role in highlighting constituents that are rhythmically or conceptually salient, such as by marking stressed syllables in words and accented words in phrases in stress languages. This feature, known as *prominence*, is a crucial function of prosody. In prosodic typology, prominence relates to how saliency is expressed through word prosody (i.e., the realization of prominence relations within words) and phrase-level prosody [i.e., the realization of prominence relations among words, cf. Jun (2005)]. Although there is no agreed consensus on the classification of word prosody [see Beckman (1996), van der Hulst (1999), and Fox (2002) for various proposals], the commonly recognized categories are tone, stress, and pitch-accent [see Beckman (1996), Ladd (1996), and Cruttenden (1997)]. At the phrase level, prominence can be realized in two ways (or their combination): (a) by marking the head of a prosodic unit or (b) by marking the edge of a prosodic unit (Hyman, 1978; Beckman, 1996; Beckman and Edwards, 1990; Ladd, 1996; Venditti *et al.*, 2014) [cf. Jun (2005)]. Head-prominence involves emphasizing a syllable or word through prosodic features like pitch, duration, and amplitude, and is typically realized as a phrasal pitch accent on the stressed syllable of the prominent word (Jun, 2005) [see Fletcher (2010) for a review of the phonetic correlates]. In contrast, edge-prominence involves marking the edges of the prominent prosodic unit with a phrasal tone at its beginning, end or both (Jun, 2005).

Evidence coming mainly from head-prominence languages indicates that articulatory gestures become longer, larger, and faster under prominence (Beckman *et al.*, 1992; Beckman and Edwards, 1994; Byrd and Saltzman, 2003; Cho, 2006). This spatiotemporal expansion is often referred to as *prosodic strengthening* [cf. Cho (2006)]. Cumulative research shows that focus as determined by information structure affects the degree of prosodic strengthening such that gestures become longer, larger, and faster across focus types from unaccented to narrow/contrastive focus [see Hermes *et al.* (2008), Roessig and Mücke (2019), and Mücke and Grice (2014) for German and Katsika *et al.* (2020, 2023) for American English]. Acoustic data also support this claim, showing that speakers mark narrow focus with longer duration, greater intensity, and higher F0 than broad focus (Breen *et al.*, 2010; Gussenhoven, 1983; Baumann *et al.*, 2006). These findings suggest that phrase-level prominence is realized by modifying several phonetic dimensions simultaneously, reflecting a hierarchy of focus types that goes beyond a simple separation between accented vs unaccented, or else focused vs unfocused. Thus, the hypothesis that focus structure is hierarchically organized and that phrase-level prominence arises from this structure, resulting in varying degrees of prominence reflecting focus type, gains support. However, most of the

a)Email: jiyoungljang@hanyang.ac.kr

evidence comes from head-prominence languages like English, where focus-induced prominence is already considered hierarchically structured: focus is typically marked by a pitch accent on the stressed syllable of the prominent word, giving rise to the hierarchy of unstressed → stressed → accented [see review in Fletcher (2010)]. Therefore, it remains uncertain whether, and if so, how, focus structure interacts with prosodic structure in edge-prominence languages, in which prominence is not assumed to project such hierarchical structure.

This study investigates Seoul Korean, an edge-prominence language without word prosody or phrasal pitch accents. In Seoul Korean, focus-induced prominence is known to be marked by the focused word initiating an accentual phrase (AP) or a higher phrase (Jun, 1993; Jeon and Nolan, 2017). The initial gestures of these focused APs have been found to be longer, larger, and faster than unfocused ones (Jang and Katsika, 2023). Additionally, boundaries of APs following the focused word and leading to the end of the intonational phrase (IP) are often regarded to be reduced or eliminated, a process known as *dephrasing* (Jun, 1993). This study primarily seeks to determine whether different types of focus result in kinematic distinctions in Seoul Korean, and if a focus hierarchy similar to that found in head-prominence languages emerges, where contrastive focus shows the strongest phonetic effect and unfocused conditions the weakest. A second goal of this study is to examine whether dephrasing is kinematically marked, and if yes, whether it varies in degree reflecting the type of focus that caused it. In our analysis, we consider the relationship between kinematic dimensions and the dynamical systems that may underlie the modulation associated with prominence, as motivated by Jang and Katsika (2024). Previous studies have shown that prosodic-structural modulations influence established kinematic relationships, namely: (a) duration increases as stiffness decreases (with stiffness measured as either normalized peak velocity over displacement or time-to-peak velocity) and (b) displacement increases as peak velocity rises [e.g., Byrd and Saltzman (1998) and Munhall *et al.* (1985)]. Of particular interest is to explore whether these relationships are further modulated by the distinct focus types emerging from information structure.

In this study, we adopt a broad view of focus as a component of information structure that highlights the informative or contrastive part of a sentence. We distinguish among focus types—broad focus, narrow focus, and contrastive focus [cf. Breen *et al.* (2010) and Gussenhoven (1983)]. These focus types are thought to induce varying degrees of prosodic prominence, which we define as the relative perceptual and articulatory salience of linguistic units, often reflected in acoustic dimensions, such as duration and pitch, and in articulatory parameters, such as gestural duration, displacement, and velocity [e.g., Hermes *et al.* (2008), Roessig and Mücke (2019), and Mücke and Grice (2014)]. We specifically examine whether these abstract focus types are kinematically encoded in Seoul Korean, meaning how they are implemented articulatorily as operationalized through gestural duration, displacement, and peak velocity. In the following, we introduce how prosodic prominence has been found to be phonetically realized in previous research, particularly through acoustic and articulatory dimensions, such as duration, displacement, and peak velocity.

A. Phonetics of prominence

Linguistic units under prominence (such as stress or accent) exhibit longer duration, higher intensity, greater pitch variation, and more distinct vowel quality in the acoustic domain (Cho, 2011; Katsika and Tsai, 2021) [cf. Lehiste (1970) and Beckman (1996); see Fletcher (2010), for an overview]. In the articulatory domain, prominence is associated with increased respiratory effort [(Ladefoged, 1967; Ladefoged and Loeb, 2002); see Lehiste (1970) for discussion]. Additionally, constriction gestures associated with prominence are characterized by longer duration, greater spatial displacement, and faster movement (Beckman and Edwards, 1994; Beckman *et al.*, 1992; Cho, 2006; de Jong, 1991, 1995; Fowler, 1995), although not all three kinematic dimensions (duration, displacement, velocity) may always be found to be employed simultaneously to distinguish between accented and unaccented gestures. For example, Beckman *et al.* (1992) found that accented syllables in English involved larger and longer jaw opening gestures compared to unaccented ones, but there was no significant difference in peak velocity. Similarly, in Cho (2006), English consonantal lip opening movements consistently showed the three kinematic patterns, but the lip closing movements were context-dependent: following an accented vowel, the movement was longer, larger, and faster, whereas preceding an accented vowel, the movement was larger but not faster or longer. While these variations may arise due to different sensitivity of the measurements taken, they may also arise from cross-linguistic differences in prominence marking. In conclusion, the dynamical mechanism that gives rise to the observed patterns of phonetic prominence is yet to be discovered.

There is furthermore evidence that prominence-related modifications do not simply mark the presence or absence of focus, but instead encode focus types as these are derived from information structure (Baumann *et al.*, 2006; Breen *et al.*, 2010; Gussenhoven, 1983; Hermes *et al.*, 2008; Katsika *et al.*, 2020, 2023; Kopera and Grigos, 2025; Mücke and Grice, 2014; Roessig and Mücke, 2019; Roessig *et al.*, 2022). For instance, acoustically, narrow focus shows longer duration, greater intensity, and higher F0 than bold focus (Breen *et al.*, 2010). Similarly, articulatory research has found that F0 and kinematic duration, displacement, and velocity increase from unfocused, to broad, to narrow and finally to contrastive focus. While there may be typological or language-specific variations, these findings indicate that phrase-level prominence is realized by adjusting multiple phonetic dimensions simultaneously and reflects distinctions beyond a simple accented vs unaccented classification [see Roessig *et al.* (2022)]. Given that, across the languages studied, phonetic

https://doi.org/10.1121/10.0039347

measures reflect the same order of focus types (contrastive focus>narrow focus>broad focus>unfocused), it is hypothesized that focus structure, as well as the phrase-level prominence that emerges from it, is organized hierarchically, with the levels of this hierarchy representing the different focus types. However, as mentioned above, most of this research comes from head-prominence languages, such as English, where phrase-level prominence is marked by a pitch accent on the stressed syllable of the prominent word [cf. Beckman and Pierrehumbert (1986)]. As a result, it remains unclear which focus types are encoded, if any, and how they are hierarchically ordered in languages with different prosodic systems, such as those that employ edge-prominence. Ultimately, addressing this question will allow a window into the basic, universal, structure of the phonological component of grammar.

B. Head-prominence vs edge-prominence

This study adopts the typological distinctions proposed by Jun (2005, 2014), as outlined in this section. In the Autosegmental-Metrical (AM) framework, prosody is described from two perspectives: the structure of an utterance and the prominence relationships within that structure [cf. Beckman (1996), Ladd (1996), and Shattuck-Hufnagel and Turk (1996)]. Prosodic structure is organized hierarchically, with higher-level prosodic units (e.g., IP, intermediate IP) consisting of one or more lower-level units [e.g., syllable, prosodic word (PWd)]. As shown in Fig. 1, syllables constitute PWd; PWd's constitute intermediate phrases (ip); and ip's constitute IPs. Prominence is marked at different levels within this structure. At the word level, certain syllables are more prominent than others, and at the phrase level, some words are more prominent than others. The realization of this structure and prominence is expressed through prosodic features such as pitch, duration, and amplitude, as well as the segmental properties of consonants and vowels.

At the word level, languages are categorized as stress (e.g., English, Greek), pitch-accent (e.g., Japanese, Basque), or tone languages (e.g., Mandarin, Hausa). Some languages combine multiple word prosody systems (e.g., both stress and lexical pitch accent, as in Chickasaw and Swedish), while others (e.g., Seoul Korean) do not use any (Jun, 2005). At the phrasal level, Jun (2005) proposes two types of prominence systems. Phrase-level prominence can be achieved by marking peaks or heads, where local adjustments to prosodic features such as pitch, duration, or amplitude make a syllable or word stand out. Languages using this system are called head-prominence languages. In the AM model, a phrase-level head is represented by a pitch accent, denoted with an asterisk (T^*, where T stands for tone, e.g., L^*, H^*). The realization of these pitch accents is at least partially shaped by the prosodic features used in the language's word-level prosody (Jun, 2005). For instance, in English, where lexical stress is marked by both duration and amplitude, the phrasal pitch accent similarly involves changes in these features. However, the specific acoustic or articulatory correlates of stress and/or phrasal pitch accents may vary between languages [(Jun, 2005); see also Cho (2011, 2015)]. It is worth also mentioning that "spillover" effects have been reported in these head-prominence languages, where an unaccented syllable following a pitch-accented syllable undergoes stress-induced strengthening (Dimitrova and Turk, 2012; Katsika and Tsai, 2021).

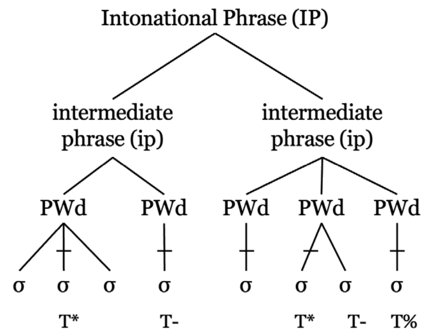

FIG. 1. Prosodic structure, adapted from Beckman and Pierrehumbert (1986). Bars between the PWd and syllable tier indicate stressed syllables as in Keating and Shattuck-Hufnagel (2002). Pitch accents are represented with T^*, phrase accents with T-, and boundary tones with T%.

In contrast, some languages realize phrase-level prominence by marking edges, where a prosodic boundary is implemented at the beginning or end of the prominent linguistic constituent, with the phrase level of that boundary (e.g., AP or ip, etc.) being language-specific. These languages are known as edge-prominence languages, and the edges of the prominent unit are marked by boundary tones. Examples of boundary tones marking prominence at various prosodic levels include the PWd boundary tone in Serbo-Croatian (%L), the phonological phrase boundary tone in Bininj Gun-wok (Lp), and the AP boundary tone in Seoul Korean (Ha) (Jun, 2005). For instance, when a word receives contrastive focus in Seoul Korean, a prosodic boundary is assumed to be inserted before the focused word, which tends to dephrase the post-focal items (see Jun, 1993, 2005; Jun and Lee, 1998). Some languages use both head- and edge-prominence strategies, and thus, these are referred to as head/edge-prominence languages (Jun, 2014).

C. Korean prosody

In this study, we investigate the articulatory correlates of focus-induced prominence in Seoul Korean, an edge-prominence language. Examining such a language offers valuable insight into how the two functions of prosody—prominence and grouping—interact. In edge-prominence languages, these functions converge in focus-marking, where grouping (through prosodic boundaries) is used to signal prominence. Additionally, as mentioned above, Seoul Korean lacks word-level prosody entirely, meaning it has no lexical stress, tone, or pitch accent (Jun, 2005). The absence of word-level prominence offers an unbiased landscape to examine the kinematics of phrase-level prominence.

Figure 2 shows the intonation model of Seoul Korean. Phrase-level prominence has been claimed to be conveyed through APs, the fundamental intonational unit, which are marked by a characteristic pitch contour (Jun, 1998). The proposed tonal pattern for APs is THLH, with the realization of the initial tone (T) influenced by the laryngeal configuration of the segment at the beginning of the AP (Jun, 1993).[1] Focus-induced prominence in Korean operates at this AP level: the focused word consistently starts an AP (or a higher-level phrase), and the subsequent AP boundaries—up to the end of the IP—are typically known to be eliminated, or possibly attenuated, a process referred to as dephrasing (Jun, 1993) (Fig. 3). Thus, APs in Seoul Korean serve dual roles: they act as the basic intonational unit and mark phrase-level prominence.

However, research on the phonetic properties of these prosodic landmarks remains relatively limited. Existing studies on the articulatory correlates of prominence in Seoul Korean have found that focused elements tend to be produced with longer durations, greater spatial displacement, and slightly faster vocalic movements [(Shin *et al.*, 2015); see Cho (2022) for a review]. More recent work investigating the contrast between "given" and "new" information has shown that broad focus—elicited in "new" information contexts—primarily influences displacement and peak velocity, with minimal effects on temporal expansion (Kim *et al.*, 2024). However, the articulatory effects of different focus types, as well as the phenomenon of dephrasing that often accompanies focus, remain understudied and focus structure overall is not well understood.

D. The current study

The first, and the main, aim of this study is to investigate whether different types of focus have distinct effects on the kinematics of the focused linguistic unit in Seoul Korean. Specifically, we explore whether prosodic realization reflects simply the presence or absence of focus or whether, similarly to head-prominence languages, varying types of focus produce fine-grained effects on articulation in Seoul Korean, a language without prosodic heads or phrase-level pitch accents. Using electromagnetic articulography, we examine how different focus types influence articulation to better understand the degrees, and, ultimately, the hierarchy of prominence in Seoul Korean. We hypothesize that prominence will be hierarchically organized, although it remains uncertain whether all focus types will be clearly differentiated. However, we expect that any observed degrees of prominence will still conform to the established hierarchy of focus types in head-prominence languages—namely, contrastive>narrow>broad>unfocused. For instance, suppose Seoul Korean uses kinematic cues to mark narrow and contrastive focus, but not broad focus. Even in this case, we would expect contrastive focus to exhibit stronger kinematic effects than narrow focus (i.e., contrastive>narrow>unfocused), and not the reverse (*narrow>contrastive>unfocused).

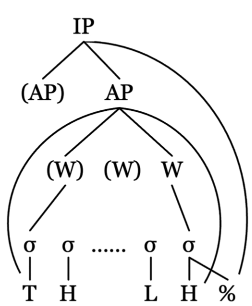

FIG. 2. Intonation model of Seoul Korean, adapted from Jun (2000).

Furthermore, as discussed at the outset of this paper, we aim to explore the relationship between kinematic parameters in order to investigate in greater depth the dynamical models that could give rise to the observed articulatory patterns. Current interdependencies between displacement and velocity on the one hand and duration and stiffness on the other hand are accounted for by critically damped second-order linear gestural dynamical systems [i.e., where duration increases as stiffness decreases, and displacement correlates with velocity, e.g., Byrd and Saltzman (1998) and Munhall *et al.* (1985)]. However, it is unclear how these interdependencies are affected by prominence, and how linear gestural dynamical systems could account for these effects.

The third question of the current study examines the kinematics of dephrasing. As discussed above, AP boundaries typically align with PWd boundaries in Korean. These boundaries are primarily marked by pitch, but phrase-initial strengthening has also been observed on the initial segment of the AP [e.g., Cho and Keating (2001)]. When AP boundaries are used for marking focus, the phenomenon of dephrasing is further observed: the last AP tones (i.e., LH rise at the beginning of the AP) of the utterance align with the onset of the PWd beginning the focused linguistic unit, and typically, the following PWd's do not bear AP tones. This pattern of pitch-related suppression that results from dephrasing is well-established. However, it is unclear whether dephrasing is also encoded kinematically. If the latter were the case, we would expect gestures at the onset of post-focal PWd's (i.e., where the tonal effects of dephrasing are observed) to be shorter, less displaced, and slower than their AP-initial counterparts. Alternatively, dephrasing of AP boundaries might only manifest in the tonal domain, with no noticeable kinematic differences. Additionally, similarly to head-prominence languages, we might potentially observe spillover effects, where the influence of focus extends to the post-focal word. (Here, we use the term "spillover" effect to refer to the influence of focus extending across a word boundary.)

Extending the third question, we examine the possibility of different focus types exerting distinct degrees/patterns of dephrasing. Based on the assumptions that (1) the effect of focus diminishes with increasing distance from the focused element and (2) that strength of focus (determined by information structure/type of focus) is reflected on the degree and scope of the focus effect, we test whether different types of focus exert

https://doi.org/10.1121/10.0039347

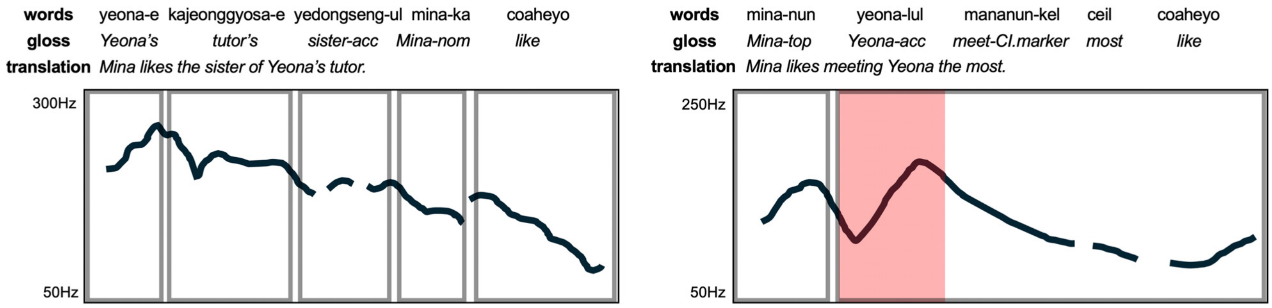

FIG. 3. Schematized pitch tracking examples from Jun (2011). The example on the left shows an utterance of broad focus with multiple APs marked by line boxes demonstrating no dephrasing. The example on the right shows dephrasing after the word bearing narrow focus marked by the shaded box. If the sentence on the right was produced under broad focus, then AP boundaries would prototypically be expected as follows: "mina-nun # yeona-lul # mananun-kel # ceil # coaheyo (# = AP boundary)."

different degrees of dephrasing and examine patterns in the scope of dephrasing by varying the length of the focused word.

Finally, as we do with the effect of focus in the second question, our fourth question investigates the relationships between the kinematic parameters on the dephrased gestures, in order to discuss the dynamic systems-based mechanism behind dephrasing.

II. METHODS

A. Participants

Six native speakers of Seoul Korean (1 female, 5 male) between the ages of 29 and 33 (mean age = 30.8; median age = 31) participated in this study. All participants were affiliated with the University of California, Santa Barbara, as graduate students or postdoctoral researchers at the time of the experiment. None of the speakers reported any speech, hearing, or vision impairments. They were unaware of the specific purpose of the study and were financially compensated for their participation.

B. Experimental procedure

Kinematic data were recorded using the AG501 3D electromagnetic articulograph (Carstens Medizinelektronik) at the UCSB SPArK (Speech, Prosody, and Articulatory Kinematics) Laboratory. Ten receiver coils were attached to various locations on each participant's head and vocal tract, including the tongue dorsum, tongue tip, a mid-point between the tongue dorsum and tongue tip, the lower and upper lips, the lower and upper incisors, the left and right ears, and the nose. The last four sensors served as reference points. Simultaneous audio recordings were captured using a Sennheiser shotgun microphone, positioned one foot from the participant's face, with a sampling rate of 16 kHz. Speech materials were displayed on a computer screen placed about three feet from the participant, using custom software, MARTA, developed by Mark Tiede (Haskins Laboratories). The collected articulatory data were smoothed and corrected for head movement based on the reference sensors, with trajectories rotated to align the X- and Y-axis to the participants' occlusal plane. To elicit appropriate focus placement in the speech material, a prompt question preceded each target utterance by one second. This timing was based on pilot recordings, in which we found a one-second interval to be sufficient for participants to plan and produce natural, fluent responses with clear focus marking. Participants read the prompt sentences silently, followed by reading the target sentences aloud.

C. Experimental design and stimuli

The test words used in the study consisted of either three (two-syllable noun + case marker /-lul/) or five syllables (four-syllable noun + case marker /-lul/) as shown in Table I. To increase the generalizability of the results across places of articulation, target words began with consonants representing three places of articulations: /m, n/, or /k/. To examine the effect of different focus types on articulation, the test words were embedded in frame sentences designed to elicit the following five focus conditions: contrastive focus (CF), narrow focus (NF), broad focus (BF), unfocused with a contrastive focus on an earlier word (UC), and unfocused with a narrow focus on an earlier word (UN). Example sentences for each focus types are presented in Table II.

To examine the kinematics of dephrasing, the word immediately following the target word was controlled to include /p/ as the onset consonant. Under the assumption that post-focal AP boundaries are subject to dephrasing through the end of the IP (Jun, 1993), the post-target word is expected to be AP-medial in all conditions except for the BF condition, where it is expected to be AP-initial.

Eight randomized repetitions of each condition were collected. The acquired data were checked for their prosodic rendition, which confirmed tonal attenuation attributed to dephrasing.

D. Measurements

Consonant (C) gesture of the initial syllable in the target word and the post-target word was analyzed using semi-automatic custom software, MVIEW, developed by Mark Tiede (Haskins Laboratories). For /m/, /n/, and /k/ of the target word, the corresponding constriction gestures were identified using the lip aperture (i.e., the distance between the upper and lower lip sensors), tongue tip vertical position,

https://doi.org/10.1121/10.0039347

TABLE I. List of target words.

List of target words		
/mapu-lul/ ("horseman-ACC")	/napi-lul/ ("butterfly-ACC")	/katɕi-lul/ ("eggplant-ACC")
/ma.tɨ.ri.tɨ-lul/ ("Madrid-ACC")	/na.mu.nul.po-lul/ ("sloth-ACC")	/ka.tɕi.tɕʰi.ki-lul/ ("pruning-ACC")

and tongue dorsum vertical position trajectories, respectively. For /p/ of the post-target word, consonant constrictors were detected on the lip aperture. The labeling procedure detected the following kinematic timepoints in each C gesture: onset, peak velocity, target, constriction maximum, release, and offset (Fig. 4). These timepoints were identified on the basis of velocity criteria, i.e., peak velocity for the homonymous timepoints, velocity minima for constriction maxima, and velocity plateaus for the other timepoints. Velocity plateaus were detected based on a set threshold of 20% of the velocity range between two consecutive alternating velocity extrema (i.e., one minimum and one maximum). This threshold, set as the default in MVIEW's labeling procedure, was chosen based on empirical observations showing it effectively approximates movement onset while conservatively avoiding "spurious onsets due to the elastic recoil of the tissue during continuous movements" [Gracco (1988), p. 4629]. On the basis of these timepoints, the following measures were calculated for each C gesture on the respective articulatory trajectories, as discussed above:

(i) Duration of gestural formation [in ms; in Fig. 4(a)]
(ii) Gestural displacement to target (i.e., the spatial difference between onset and maximum constrictor position) [in mm; in Fig. 4(b)]
(iii) Gestural peak velocity to target [in cm/s; in Fig. 4(c)]

TABLE II. Example stimulus sentences by Focus type (CF, NF, BF, UC, and UN) with the test word /mapu-lul/ ("horseman-ACC"). Test words are underlined and focused words are in bold.

Focus	Example sentences
CF	Prompt sentence: "Did Minam visit the farmer?" Test sentence: minamika **mapulul** paŋmunhesʌ Minam-NOM **horseman**-ACC visit-PAST "Minam visited **the horseman**."
NF	Prompt sentence: "Who did Minam visit?" Test sentence: minamika **mapulul** paŋmunhesʌ Minam-NOM **horseman**-ACC visit-PAST "Minam visited **the horseman**."
BF	Prompt sentence: "What happened?" Test sentence: **minamika** **mapulul** **paŋmunhesʌ** **Minam-NOM** **horseman-ACC** **visit-PAST** "**Minam visited the horseman**."
UC	Prompt sentence: "Did Junseok visit the horseman?" Test sentence: **minam**ika mapulul paŋmunhesʌ **Minam**-NOM horseman-ACC visit-PAST "**Minam** visited the horseman."
UN	Prompt: sentence "Who visited the horseman?" Test sentence: **minam**ika mapulul paŋmunhesʌ **Minam**-NOM horseman-ACC visit-PAST "**Minam** visited the horseman."

E. Statistical analysis

The retrieved data were analyzed by linear mixed-effects analysis using the *lme4* package (Bates *et al.*, 2015) in R (R Core Team, 2024). To test the effect of focus type, initial C gesture of the target word was examined for the dependent variables of formation duration, displacement, and peak velocity. Fixed effects of *focus type* (CF, NF, BF, UC, UN) and *word length* (3-syllable, 5-syllable) were added. Initial C gesture of the post-target word was also analyzed to investigate the effect of *focus type* and *word length* on the post-target word and the pattern of dephrasing. For both analyses, random effects of *speaker* as well as *place of articulation* (bilabial, alveolar, velar) were added. (Note that since the goal of the present study is to enhance the generalizability of the findings across consonants with different places of articulation—rather than to examine how each articulatory trajectory behaves—*place of articulation* was included as a random factor instead of as a fixed factor in the statistical models.)

For the analyses on both the target and the post-target words, a second set of linear mixed-effects models were built to examine the relationship between kinematic parameters as a function of different focus types. To better understand how focus-induced prominence influences individual kinematic parameters (e.g., stiffness, peak velocity, displacement), we employed statistical models that account for the known covariation among these measures, applying the methodology in Katsika and Tsai (2021). As established in previous articulatory studies [e.g., cf. Edwards *et al.* (1991) and Byrd and Saltzman (2003)], kinematic parameters such as duration, stiffness, peak velocity, and displacement are inherently interrelated. For example, stiffness and duration typically show an inverse relationship—stiffer gestures tend to have shorter durations—while displacement and peak velocity are positively correlated, with larger gestures generally exhibiting higher peak velocities.

To examine the effects of prosodic structure on a given parameter while controlling for its relationship with other variables, we included relevant kinematic covariates in our

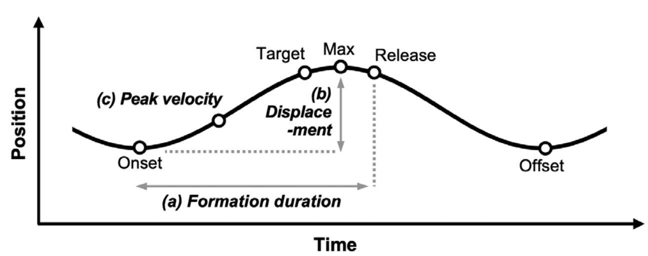

FIG. 4. Schematized constriction gesture with kinematic measurements.

https://doi.org/10.1121/10.0039347

models as fixed factors. Specifically: (a) when analyzing duration, we included peak velocity normalized over displacement, referred to as stiffness henceforth for reasons of brevity (see below for justification), as a continuous control factor[2]; (b) when analyzing displacement, we included peak velocity as a control factor; (c) when analyzing peak velocity, we included displacement as a control factor. Note that the ratio of peak velocity to displacement has been proposed as an empirical estimate of kinematic stiffness, which captures the observation that peak velocity varies with displacement and has been shown to increase as movement duration decreases in research that views gestures as critically damped second-order linear systems (Munhall *et al.*, 1985; Ostry and Munhall, 1985); used in e.g., Beckman *et al.* (1992), Hawkins (1992), and Roon *et al.* (2007). This modeling strategy allowed us to assess whether prosodic conditions (e.g., focus type) systematically influenced each kinematic parameter, beyond what could be explained by interdependencies among articulatory measures. In doing so, we aim to provide a more precise characterization of how prosodic strengthening shapes specific aspects of gestural dynamics. Pairwise comparisons were assessed by the *emmeans* package (Lenth, 2023) with Tukey adjustment, and the results are provided in the Appendix.

III. RESULTS

A. Initial C gesture of the target word

A main effect of *Focus type* was significant in all three measured dimensions, i.e., formation duration [$F(4, 1456) = 93.2$, $p < 0.001$], displacement [$F(4, 1456) = 16.5$, $p < 0.001$], and peak velocity [$F(4, 1456) = 7.0$, $p < 0.001$]. Figure 5 plots the results on formation duration, displacement, and peak velocity, respectively (see Table III in Appendix for statistics on pairwise comparisons). Kinematic dimensions appear to differentiate among focus types and not simply marking presence vs absence of focus. However, dimensions differed in the number of focus types, and thus degrees of prominence they distinguished. Formation duration presented three degrees of prominence: CF, NF > BF > UN, UC [Fig. 5(a)]. Displacement showed two degrees: CF, NF > BF, UN, UC [Fig. 5(b)]. Peak velocity also distinguished two degrees, but with BF not being significantly different from either CF/NF or UN/UC: CF, NF, BF > BF, UN, UC [Fig. 5(c)]. These effects held regardless of the number of syllables in the target word, since *focus type* did not interact with *word length* in any of the measures ($p > 0.05$, in all measurements).

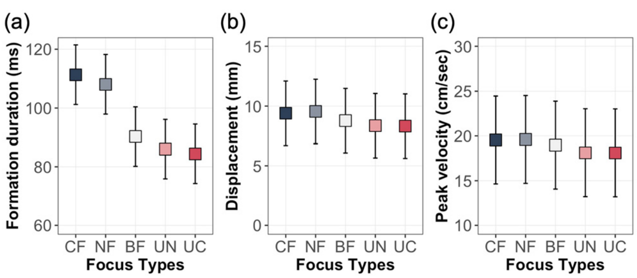

FIG. 5. Fitted results on (a) formation duration (in ms), (b) displacement (in mm), and (c) peak velocity (in cm/s) for target word initial C gesture.

In short, all three articulatory measures did not simply encode focused vs unfocused distinction but instead differentiated levels that reflect the focus structure derived from information structure. Articulatory duration distinguished the most levels (i.e., three). Length of the focused word did not affect the distinction of levels.

B. Relationships between kinematic parameters on initial C gesture of the target word

Readers are reminded that a second set of linear mixed-effects models were fitted to understand the relationship between kinematic dimensions and the dynamical systems that might account for the articulatory modulation of prominence. The upper panel of Fig. 6 plots the fitted results at average value of the continuous factors (i.e., stiffness for formation duration, peak velocity for displacement, and displacement for peak velocity) by focus type. The fitted relationships between each measurement and continuous factor are plotted by focus type in the lower panel of Fig. 6. As predicted, formation duration and stiffness (i.e., normalized peak velocity over displacement) showed an inverse relationship such that formation duration increased with decrease in stiffness [$F(1, 1434) = 709.7$, $p < 0.001$]. For formation duration, a main effect of focus types was significant [$F(4, 1446) = 49.5$, $p < 0.001$]. In this analysis, formation duration distinguished more levels of prominence compared to the one fitted in Sec. III A. Specifically, formation duration in CF was significantly longer than that in NF, presenting four degrees of prominence: CF > NF > BF > UN, UC [Fig. 6(a); see Table IV in the Appendix for

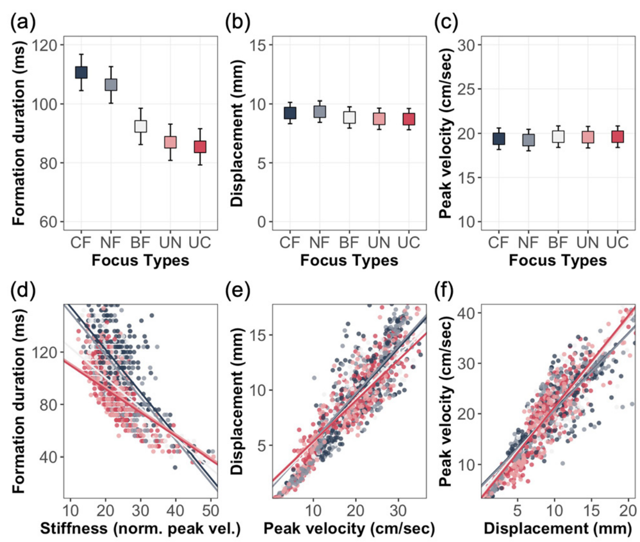

FIG. 6. Top panel: Fitted results on (a) formation duration at average stiffness (in ms), (b) displacement at average peak velocity (in mm), and (c) peak velocity at average displacement (in cm/s) for target word initial C gesture. Bottom panel: Fitted regression lines for (d) formation duration and stiffness (i.e., normalized peak velocity), (e) displacement and peak velocity, and (f) peak velocity and displacement as a function of focus types.

statistics on pairwise comparisons]. A two-way interaction of focus types and stiffness was significant [$F(4, 1446) = 18.9$, $p < 0.001$]. As illustrated in Fig. 6(d), pairwise comparisons on fitted regressions revealed that the slope values of the regressions were significantly different between the following two groups: CF, NF vs BF, UN, UC (see Table V in the Appendix for statistics on pairwise comparisons). In particular, the slopes of the regressions in CF and NF are significantly steeper than that in BF, UN, and UC. That is, given the same decrease in stiffness, the increase in duration is greater when the gesture is under CF or NF than when it is under BF or out-of-focus.

For displacement, there was a positive relationship with peak velocity such that displacement increased with increase in peak velocity, as expected [$F(1, 1439) = 2252.8$, $p < 0.001$]. While the main effect of focus type was not significant, there was a significant interaction between focus type and peak velocity [$F(1, 1446) = 7.9$, $p < 0.001$]. At average peak velocity, displacement distinguished two levels of prominence: CF, NF > BF, UN, UC, similar to the analysis in Sec. III A [Fig. 6(b); see Table IV in the Appendix for statistics on pairwise comparisons]. In terms of the slopes of the regressions, two degrees were distinguished: CF, NF, BF > BF, UN, UC, with BF not being significantly different from either CF/NF or UN/UC (see Table V in the Appendix for statistics on pairwise comparisons).

Finally, peak velocity, in general, shows an expected positive relationship with displacement such that peak velocity increases with increase in displacement [$F(1, 1127) = 2252.8$, $p < 0.01$]. While a main effect of focus types was significant [$F(4, 1444) = 10.8$, $p < 0.001$], pairwise comparisons at average displacement revealed no significant difference between the five levels [Fig. 6(c)]. However, a significant two-way interaction of focus type and displacement [$F(4, 1445) = 17.6$, $p < 0.001$] suggested that the slope values of the regressions were significantly different between the following two groups: CF, NF vs BF, UN, UC (see Table V in the Appendix for statistics on pairwise comparisons). Interestingly, the slopes for BF, UN, and UC group were higher than that of CF and NF group, indicating that given the same increase in displacement, the increase in peak velocity is less in CF, NF than in BF, UN, and UC. The pattern of shallower slope of peak velocity against displacement has been reported as a signature of boundary-related modulation in Seoul Korean in Jang and Katsika (2024). More on this will be addressed in Sec. IV.

C. Interim discussion on the kinematics of focus marking

This section summarizes the results on the initial C gesture of the target word. Results in Sec. III A suggest that prominence in Seoul Korean does encodes more levels beyond the simple focused vs unfocused distinction. Specifically, formation duration distinguished three degrees of prominence: CF, NF > BF > UN, UC. Displacement and peak velocity both showed two levels with BF being grouped with the unfocused conditions for displacement, and BF not being significantly distinguished with either group for peak velocity.

When we consider the kinematic relationships (Sec. III B), results not only show expected co-dependencies between parameters but also confirm that these relationships are indeed affected by prosodic-structural modulations. Across all three measurements, the regression slopes were consistently distinguished in two levels—(1) CF and NF in one level and (2) BF, UN, and UC in another level (although BF was not significantly different from either group for displacement). Another interesting finding is that results considering the co-dependency between duration and stiffness provide more fine-grained distinctions among focus types. That is, while CF and NF did not differentiate each other on the surface level (Sec. III A), analysis with stiffness provides another level distinguishing CF to NF. Finally, when taking the kinematic relationships into account, the focus effect observed in peak velocity in Sec. III A becomes nonsignificant, indicating that the previously observed effect was likely driven by variations in gesture displacement. Moreover, the shallower slope of peak velocity found for more prominent focus conditions resemble the kinematic signature found for boundary-related modulation in Seoul Korean (Jang and Katsika, 2024), and will be discussed in Sec. IV.

D. Initial C gesture of the post-target word

Results on the initial C gesture of the post-target word are plotted in Fig. 7. For formation duration, while there was no main effect of *focus type*, there was a significant interaction of *focus type* with *word length* [$F(4, 1482) = 6.2$, $p < 0.001$]. *Post hoc* analysis revealed that the comparisons among focus types on the 3-syllable target words did not show a significant difference in the formation duration of the initial C gesture in the post-target word, except for the comparison between NF and UC. Formation duration for NF was longer than that for UC ($\beta = 3.8$, $p < 0.05$). However, when the target word was five syllables long, the post-target C gesture was *shorter* in CF as compared to either of the unfocused conditions (CF-UN: $\beta = -4.3$, $p < 0.01$; CF-UC: $\beta = -3.7$, $p < 0.05$) (Fig. 8). Other focus types were not systematically distinguished.

As for displacement, there was a significant main effect of *focus type* [$F(4, 1482) = 5.2$, $p < 0.001$]. Pairwise comparison indicated that post-target word-initial C was *more*

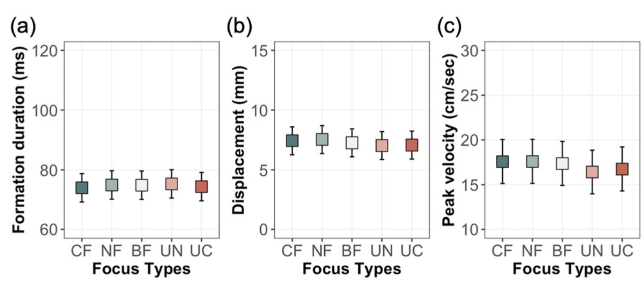

FIG. 7. Fitted results on (a) formation duration (in ms), (b) displacement (in mm), and (c) peak velocity (in cm/s) for post-target word initial C gesture.

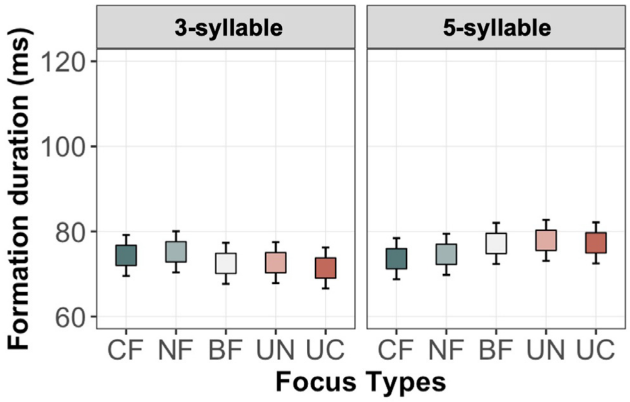

FIG. 8. Fitted results on formation duration (in ms) for post-target word initial C gesture by *word length* (3-syllable, 5-syllable).

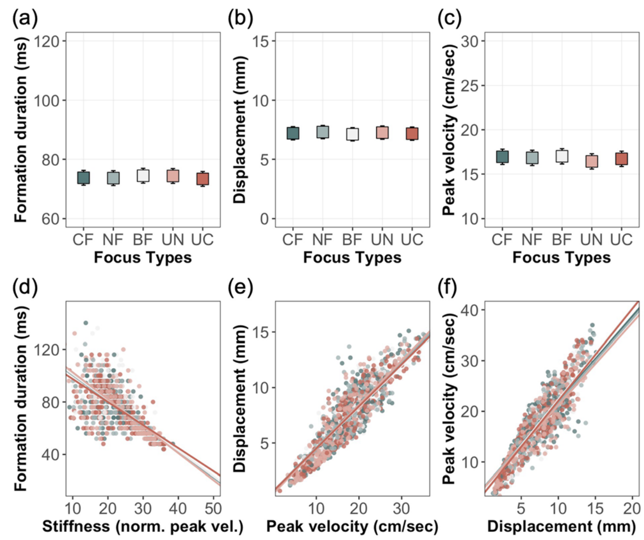

FIG. 9. Top panel: Fitted results on (a) formation duration at average stiffness (in ms), (b) displacement at average peak velocity (in mm), and (c) peak velocity at average displacement (in cm/s) for post-target word initial C gesture. Bottom panel: Fitted regression lines for (d) formation duration and stiffness (i.e., normalized peak velocity), (e) displacement and peak velocity, and (f) peak velocity and displacement as a function of focus types.

displaced in CF and NF as compared to UN ($\beta = 0.4$, $p < 0.05$; $\beta = 0.5$, $p < 0.01$, respectively) and UC ($\beta = 0.5$, $p < 0.01$; $\beta = 0.4$, $0.05 < p < 0.07$, respectively) [Fig. 7(b)]. Thus, the focus effect on displacement seems to spill over to the following word. *Focus type* did not further interact with *word length*.

In the peak velocity measure, a significant main effect of *focus type* was also detected [$F(4,\ 1482) = 6.9$, $p < 0.001$]. *Post hoc* analysis showed that post-target word-initial C was faster in CF and NF than in UN [$\beta = 1.2$, $p < 0.001$; $\beta = 1.2$, $p < 0.001$, respectively] and UC [$\beta = 0.8$, $p < 0.05$; $\beta = 0.8$, $p < 0.05$, respectively], again presenting a pattern that suggests spillover effects of focus on the post-focal word [Fig. 7(c)]. Also, the gesture was faster in BF than in UN ($\beta = 1.0$, $p < 0.01$). *Focus type* did not further interact with *word length*.

In sum, spillover of the focus effect as well as post-focal compression are observed on the initial C gesture of the post-target word. Specifically, spillover effects are observed on displacement and peak velocity measures, such that gestures following more prominent focused elements show larger and faster movements. On the other hand, for formation duration, possible post-focal temporal compression is found when the target word is long and under CF.

E. Relationships between kinematic parameters on initial C gesture of the post-target word

Similar to the analysis of the initial C gesture in the target word, a second set of linear mixed-effects models that account for the relationship between kinematic dimensions was fitted for the initial C gesture in the post-target word. Top panel of Fig. 9 plots the fitted results at average value of the continuous factors (i.e., stiffness for formation duration, peak velocity for displacement, and displacement for peak velocity) by focus type. The fitted relationships between each measurement and continuous factor are plotted by focus type in the bottom panel of Fig. 9. As predicted, all continuous factors showed significant effects in the expected direction: (1) inverse relationship of stiffness (i.e., normalized peak velocity over displacement) with formation duration [$F(1,\ 1143) = 681.8$, $p < 0.001$]; (2) positive relationship of peak velocity with displacement [$F(1,\ 1440) = 2896.4$, $p < 0.001$]; and (3) positive relationship of displacement with peak velocity [$F(1,\ 1454) = 2896.9$, $p < 0.001$].

For formation duration, while there was no main effect of focus type, in this analysis, a significant focus type × word length interaction was detected [$F(1,\ 1472) = 2.6$, $p < 0.05$]. This interaction could have stemmed from the marginal effect of NF being shorter than UN only in 5-syllable word condition [$\beta = -2.8$, $0.05 < p < 0.08$], possibly showing temporal compression under NF. However, pairwise comparisons indicated no significant difference among the focus types in both 3-syllable and 5-syllable word conditions. As for displacement, no significant main effect nor interaction was found related to focus types. Finally, for peak velocity, there was a main effect of focus type [$F(1,\ 1472) = 3.4$, $p < 0.01$] as well as a significant interaction between focus type and displacement [$F(1,\ 1472) = 4.8$, $p < 0.001$]. Pairwise comparisons revealed that, at average displacement, the initial C gesture of the post-target word in CF and BF was significantly faster than that in UN [$\beta = 0.5$, $p < 0.05$; $\beta = 0.6$, $p < 0.01$, respectively]. In terms of the fitted regressions, pairwise comparisons yielded unclear results: the slope values of the regressions in BF and UC were steeper compared to those in NF and UN (NF-BF: $\beta = -0.1$, $p < 0.05$; NF-UC: $\beta = -0.2$, $p < 0.05$; BF-UN: $\beta = 0.2$, $p < 0.05$; UN-UC: $\beta = -0.2$, $p < 0.01$).

F. Interim discussion on the kinematics of dephrasing

This section summarizes the results on the initial C gesture of the post-target word. Results in Sec. III D possibly indicate that the effect of focus, which begins in the AP-initial constriction gesture, spans over several consecutive constrictions [as also found in Jang and Katsika (2023)], and that, in the case of relatively short words, as in our 3-syllable condition here, the effect spills over to the post-focal word. In the case of 5-syllable words, however, the effect of focus ends before the full length of the word is reached, and temporal compression is instead observed in the initial C gesture of the post-focal word. Temporal compression has been documented in the literature as occurring either before or after temporal expansion, particularly in boundary-related temporal modulations. This phenomenon is often interpreted as a compensatory mechanism to offset the effects of lengthened durations (Byrd *et al.*, 2006). While further studies are warranted to confirm whether intervening gestures are affected by similar mechanisms, the fact that compression is only found in the context of CF in the current study may support the hypothesis that this type of focus exerts one of the strongest degrees of prominence.

However, results suggest that the spillover effects observed in earlier analyses in Sec. III D disappear when we consider the co-variance between kinematic parameters. This may be possibly due to some sort of equalizing effect that brings articulatory movements back to a regular oscillating cycle, counteracting the initial focus-induced enlargement and acceleration.

IV. GENERAL DISCUSSION

The present study examined whether different focus types—contrastive focus, narrow focus, broad focus, and out-of-focus—are kinematically distinguished in Seoul Korean. The patterns of the current dataset indicate that the kinematic properties of articulatory gestures in Korean do not merely reflect the presence or absence of focus; rather, they encode focus structure, displaying cumulative effects from out-of-focus to contrastive focus, similar to the patterns observed in head-prominence languages [e.g., Hermes *et al*. (2008), Mücke and Grice (2014), Roessig and Mücke (2019), Roessig *et al*., 2022 for German; see also Katsika *et al*. (2020, 2023) for American English]. Notably, different kinematic parameters revealed varying degrees of prominence, potentially due to the different granularity of their respective measurements. Among these, formation duration exhibited the greatest sensitivity to prominence, distinguishing the most levels of focus: the word-initial consonant gesture of the target word was longest under contrastive or narrow and shortest in unfocused. This finding provides support to the idea that articulatory strengthening does not simply mark AP boundaries but also encodes focus structure. Displacement distinguished two levels of prominence, differentiating between narrow/contrastive focus and broad/out-of-focus conditions (CF/NF vs BF/UC/UN). Peak velocity distinguished two levels between narrow/contrastive and unfocused conditions with broad focus not being separated from either the higher prominence group (CF/NF) or the lower prominence group (UC/UN). Despite the different number of prominence levels encoded by each kinematic dimension, the overall ordering of focus types remained consistent across dimensions, with contrastive and narrow focus at the top and the unfocused conditions at the bottom. Thus, paralleling findings from head-prominence languages, data from Seoul Korean, an edge-prominence language, suggest a universal order and hierarchical structure of focus. This offers valuable insight into how information-structural focus interacts with prosodic realization across typologically distinct languages. Katsika *et al*. (2020), in their study of American English, argue that different focus types yield distinct pitch accents and are accompanied by parallel modifications in temporal dimensions, such as articulatory duration. This alignment of tonal and temporal cues was interpreted as strong evidence that focus types correspond to distinct levels in a prominence hierarchy in English. In contrast, Seoul Korean does not vary its AP tonal shape across focus types, with previous descriptions largely pointing to a binary distinction whether the default LH rise is maintained (i.e., focused) or suppressed through dephrasing (i.e., unfocused). Nonetheless, our results show that Seoul Korean speakers produce distinct articulatory patterns even among focus types that are expected to share the same boundary structure. This suggests that in an edge-prominence system like Korean, the phonetic encoding of focus structure extends beyond the insertion or suppression of prosodic boundaries and involves more gradient modulation of articulatory parameters as a function of focus type.

As part of our second research question, we additionally analyzed the effect of focus types on the focused linguistic unit, taking into account the established relationships between kinematic parameters. Results of the current study suggest that when we do consider the interrelationship between the kinematic parameters, we see finer-grained distinctions between the focus levels in the duration measure. That is, four degrees of prominence were observed: CF > NF > BF > UN, UC, even differentiating contrastive from CF. Displacement distinguished two degrees of prominence, drawing the line between narrow and broad focus. Peak velocity, however, showed an interesting result when corresponding displacement was taken into account: No significant differences were found among the five focus types, suggesting that when displacement is controlled, the movement velocity of the gestures appears to remain unaffected by focus. This may suggest that the effects of focus on velocity might be mediated or overshadowed by the inherent relationship between displacement and velocity in the articulatory system.

As for the origin of this hierarchical structure, this remains uncertain. One possibility is that it reflects an inherent hierarchical order in information structure that interfaces with prosodic structure. Alternatively, the hierarchy may arise from varying degrees of trade-offs between linguistic and prosodic information associated with different focus types, as suggested by the Speech Signal Redundancy Hypothesis (Aylett and Turk, 2004). In this view, information structure might not be directly encoded within prosodic structure; instead, the strength of prosodic effects could be influenced by the level of linguistic redundancy of the target word under different focus conditions. In experimental settings where participants must use the same word order across different focus conditions, the linguistic redundancy of the focused word is likely to decrease from broad to narrow and from narrow to contrastive focus, requiring speakers to rely more on prosodic features to convey prominence. However, it remains to be seen whether a similar level of granularity in focus marking would be observed when speakers are free to employ other linguistic strategies, such as varied word orders or morphological markers, to indicate focus.

As current theories of prosody stand, a specific account of the connection between information structure to prosodic structure is lacking. For instance, in the framework of articulatory phonology, modulation gestures, or mu-gestures, have been proposed to model prominence-related manifestation of gestures (Saltzman *et al.*, 2008). In this model, the cumulative effect of focus type could in principle be captured by mu-gestures of different strengths, similarly to the control of pi-gesture strength that has been used to model effects of boundaries at different levels of the prosodic hierarchy [cf. Byrd and Saltzman (2003)]. However, an equivalent hierarchy of prominence in which focus type is encoded has not been established [see, for instance, hierarchies proposed in metrical phonology; e.g., Liberman and Prince (1977), Nespor and Vogel (1986), Pierrehumbert and Beckman (1988), and Selkirk (1984)]. The non-linear dynamical model of prominence marking proposed in Roessig and Mücke (2019) captures the cumulative (kinematic and F0) effects of prominence across focus types by controlling a single dynamical parameter. This model views prominence as a gradient system, predicting an increasing degree in prominence and a set of phonetic parameters marking them that work in tandem. A direct link to either information or prosodic structure levels/categories is not assumed. Why, however, these degrees would correspond to the specific focus types and in that order in languages is not clear. The results of the present study call for further development of theories on the connection between information structure and prosodic structure.

Another contribution of the current study pertains to the phenomenon of dephrasing, the kinematic dimensions of which are not well understood. Prior to analyzing the kinematic aspects, utterances were evaluated for dephrasing in the tonal domain, and the analysis confirmed that the data indeed exhibited tonal dephrasing. With the exception of some temporal compression following contrastive focus—and this only in cases of relatively long focused words, no robust evidence of articulatory attenuation is found to be accompanying the tonal attenuation of dephrasing. Specifically, we did not observe significant kinematic weakening (i.e., shorter duration, less displaced, and/or slower velocity) in the post-focal position when compared to AP-initial position, though this conclusion is based on a limited speaker sample. Instead, initial gestures of immediately post-focal words are found to be strengthened under focus. These patterns are consistent with an account of focus having spillover prosodic effects (Dimitrova and Turk, 2012; Katsika and Tsai, 2021). It is, then, proposed that the activation of the mu-gesture, possibly in-phased with the focused linguistic unit's initial constriction gesture (Jang, 2023), extends from the left-edge, reaching gestures that are several syllables away. Such effects are also detected in Jang and Katsika (2023), where it is shown that the scope of focus-induced strengthening extends beyond the initial syllable of the focused AP and reaches up to the third syllable of the AP. Although no systematic kinematic attenuation was detected on dephrased APs, based on the patterns found in the current study, it may be suggested that focus effects cross the word boundary between the focused word and the post-focal one precisely because of the dephrased AP boundary. In that sense, the spillover effects themselves might be the kinematic signature of dephrasing. The account of spillover effects of focus is further enhanced by the fact that the temporal compression after contrastive focus was constrained by the number of syllables intervening between the beginning of the focused AP and the post-focal word (i.e., the length of the focused word): In the case of short focused words, the durational effects of focus cross over the word boundary and to the post-focal word; in the case of longer focused words, the durational effect of focus, which begins with the onset of the focused word, dies out before it reaches the end of the word, which is several syllables away. Future studies will further examine this account.

Finally, the present study finds that the interrelationships between kinematic parameters, i.e., the relationship between duration and normalized peak velocity over displacement (referred to as stiffness), and the relationship between displacement and peak velocity, are further modulated by the type of focus. Especially for duration, which showed the greatest distinction between focus types, robust effects of prosodic position on the relationship between duration and stiffness (estimated by normalized peak velocity over displacement) was detected, forming two groups: (1) contrastive and narrow focus, (2) broad focus and two unfocused conditions. In terms of displacement, two groups were distinguished: (1) contrastive and narrow focus and (2) two unfocused conditions with broad focus not being significantly different from either of the groups. The direction was that gestures receiving narrow focus, regardless of whether it was contrastive or not, had steeper slopes than the broad focus or unfocused conditions. While peak velocity also distinguished two groups of (1) contrastive and narrow focus, (2) broad focus and two unfocused conditions, it showed an opposite direction such that the more prominent group had shallower slopes than the non-prominent group, showing less increase in peak velocity, given the same increase in displacement. This result of the opposite direction is interesting in the sense that the shallower slope of peak velocity against displacement was reported as a boundary-related signature of Seoul Korean (Jang and Katsika, 2024). Readers are reminded that Seoul Korean is an edge-prominence language known to mark prominence by the means of inserting an AP (or a higher phrase), placing the focused linguistic unit to be aligned with the left-edge of the phrase. It is possible that in Seoul Korean, prominence is

https://doi.org/10.1121/10.0039347

manifested through both functions of prosody—prominence and phrasing, and thus is exhibiting a trait of boundary-related marking in terms of the articulatory modulation.

Combined with the results of Jang and Katsika (2024) on the kinematic profile of IP boundaries, it may be concluded that the prosodic functions of prominence- and IP boundary-marking present distinct stiffness profiles on articulation, suggesting that such prosodic modulations might be arising at least partly from controlling gestural stiffness [see discussion in Cho (2006)]. Stiffness might not be the only parameter controlled though, since this would not be able to capture the displacement effects. For that, we would need to control the gestural target. Note, however, that the relationship between displacement and peak velocity does not present the same dependency on prosodic position as the relationship between duration and stiffness. According to Cho (2006), shrinking (or less shrinking), defined as "a change in both target and stiffness," may account for the effect found for duration and displacement, but not in peak velocity. In the mu-gestural model, we would need both a spatial and a temporal mu-gesture to be co-active in order to get the longer and larger gestures found here. However, mu-gestures would not be able to capture the distinct stiffness profiles presented by prominent and IP-final positions [see also discussion in Jang and Katsika (2024) and Iskarous and Pouplier (2022)]. On the other hand, the similarity between the patterns found here and those in Roessig and Mücke (2019), along with the distinct profiles shown by the relationship between duration and peak velocity normalized over displacement (termed stiffness) by focus type in the current study and by prosodic position in Jang and Katsika (2024), may be seen as supporting proposals to reconsider the dynamics of gestures and their prosodic modulations through non-linear dynamics [cf. Roessig and Mücke (2019), Sorensen and Gafos (2016), and Iskarous and Pouplier (2022)].

V. CONCLUSION

In conclusion, the present study demonstrates that the kinematic properties of articulatory gestures in Seoul Korean encode a hierarchical structure of focus types, ranging from out-of-focus to contrastive focus, similar to observations in head-prominence languages. Different kinematic dimensions, such as formation duration, displacement, and peak velocity, exhibit varying degrees of prominence reflecting different types of focus. Evidence that articulatory strengthening is not solely governed by the presence or absence of prominence but is also shaped by focus structure suggests that the connection between information structure and prosodic structure is more complex than previously understood. Moreover, the study's findings on post-focal gestures suggest that dephrasing may attenuate AP boundaries to the extent that focus-induced strengthening crosses over to the post-focal word. Finally, the intricate interrelationships between kinematic parameters reported in the current study as well as in Jang and Katsika (2024) call for models that account for the spatial and temporal dimensions of articulatory gestures to take into account the relationships between the parameters. Consequently, this research contributes to advancing our understanding of how focus structures are realized phonetically and offers a foundation for further exploration of the interplay between articulation, prosody, and focus structure.

ACKNOWLEDGEMENTS

The authors thank the Korean speakers who participated in the study. We are also extremely grateful to Louis Goldstein, Matthew Gordon, and John Du Bois for their comments on earlier versions of the manuscript. We thank the associate editor and the anonymous reviewers for their thoughtful comments and constructive feedback, which greatly helped improve the clarity and quality of the manuscript. This work was supported by National Science Foundation Grant No. 1551428.

AUTHOR DECLARATIONS

Conflict of Interest

The authors declare no conflicts of interest related to the research and hereby state that this project has obtained approval from the institutional review board (IRB) to ensure compliance with ethical guidelines.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX

Results of pairwise comparisons by focus type are presented in Tables III–V. The tables report comparisons for formation duration, displacement, and peak velocity, corresponding to the analyses described in Secs. III A and III B.

TABLE III. Results of pairwise comparisons by Focus Type for formation duration, displacement, and peak velocity for the initial C gesture of the target word in Sec. III A. (***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; n.s.: $p > 0.06$.)

		CF	NF	BF	UN
NF	Dur.	$\beta = 3.27$ *n.s.*			
	Displ.	$\beta = -0.14$ *n.s.*			
	Peak vel.	$\beta = -0.06$ *n.s.*			
BF	Dur.	$\beta = 21.07^{***}$	$\beta = 17.79^{***}$		
	Displ.	$\beta = 0.62^{*}$	$\beta = 0.77^{*}$		
	Peak vel.	$\beta = 0.57$ *n.s.*	$\beta = 0.63$ *n.s.*		
UN	Dur.	$\beta = 25.35^{***}$	$\beta = 22.08^{***}$	$\beta = 4.28$ *n.s.*	
	Displ.	$\beta = 1.04^{***}$	$\beta = 1.19^{***}$	$\beta = 0.42$ *n.s.*	
	Peak vel.	$\beta = 1.42^{**}$	$\beta = 1.48^{**}$	$\beta = 0.85$ *n.s.*	
UC	Dur.	$\beta = 26.94^{***}$	$\beta = 23.66^{***}$	$\beta = 5.87^{*}$	$\beta = 1.59$ *n.s.*
	Displ.	$\beta = 1.08^{***}$	$\beta = 1.23^{***}$	$\beta = 0.46$ *n.s.*	$\beta = 0.03$ *n.s.*
	Peak vel.	$\beta = 1.44^{**}$	$\beta = 1.50^{**}$	$\beta = 0.86$ *n.s.*	$\beta = 0.01$ *n.s.*

TABLE IV. Results of pairwise comparisons at average stiffness, peak velocity, or displacement by Focus Type for formation duration, displacement, and peak velocity for the initial C gesture of the target word in Sec. III B. (***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; n.s.: $p > 0.06$.)

		CF	NF	BF	UN
NF	Dur.	$\beta = 4.19^{*}$			
	Displ.	$\beta = -0.11$ *n.s.*			
	Peak vel.	$\beta = 0.14$ *n.s.*			
BF	Dur.	$\beta = 18.23^{***}$	$\beta = 14.04^{***}$		
	Displ.	$\beta = 0.37^{*}$	$\beta = 0.48^{**}$		
	Peak vel.	$\beta = -0.24$ *n.s.*	$\beta = -0.38$ *n.s.*		
UN	Dur.	$\beta = 23.62^{***}$	$\beta = 19.43^{***}$	$\beta = 5.39^{**}$	
	Displ.	$\beta = 0.49^{**}$	$\beta = 0.60^{***}$	$\beta = 0.11$ *n.s.*	
	Peak vel.	$\beta = -0.18$ *n.s.*	$\beta = -0.32$ *n.s.*	$\beta = 0.05$ *n.s.*	
UC	Dur.	$\beta = 25.16^{***}$	$\beta = 20.98^{***}$	$\beta = 6.94^{***}$	$\beta = 1.55$ *n.s.*
	Displ.	$\beta = 0.51^{***}$	$\beta = 0.62^{***}$	$\beta = 0.13$ *n.s.*	$\beta = 0.02$ *n.s.*
	Peak vel.	$\beta = -0.23$ *n.s.*	$\beta = -0.38$ *n.s.*	$\beta = 0.00$ *n.s.*	$\beta = -0.05$ *n.s.*

TABLE V. Results of pairwise comparisons of the regression slopes by Focus Type for formation duration, displacement, and peak velocity for the initial C gesture of the target word in Sec. III B. (***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; tr.: $0.05 < p < 0.06$; n.s.: $p > 0.06$.)

		CF	NF	BF	UN
NF	Dur.	$\beta = -0.02$ *n.s.*			
	Displ.	$\beta = 0.00$ *n.s.*			
	Peak vel.	$\beta = 0.00$ *n.s.*			
BF	Dur.	$\beta = -0.96^{***}$	$\beta = -0.94^{***}$		
	Displ.	$\beta = 0.02$ *n.s.*	$\beta = 0.03$ *tr.*		
	Peak vel.	$\beta = -0.20^{***}$	$\beta = -0.19^{***}$		
UN	Dur.	$\beta = -1.47^{***}$	$\beta = -1.44^{***}$	$\beta = -0.50$ *n.s.*	
	Displ.	$\beta = 0.05^{**}$	$\beta = 0.06^{***}$	$\beta = 0.02$ *n.s.*	
	Peak vel.	$\beta = -0.27^{***}$	$\beta = -0.27^{***}$	$\beta = -0.07$ *n.s.*	
UC	Dur.	$\beta = -1.46^{***}$	$\beta = -1.44^{***}$	$\beta = -0.50$ *n.s.*	$\beta = 0.00$ *n.s.*
	Displ.	$\beta = 0.05^{**}$	$\beta = 0.05^{***}$	$\beta = 0.02$ *n.s.*	$\beta = 0.00$ *n.s.*
	Peak vel.	$\beta = -0.29^{***}$	$\beta = -0.29^{***}$	$\beta = -0.09$ *n.s.*	$\beta = -0.02$ *n.s.*

[1]H tone when the onset consonant is a fortis or aspirated stop/affricate, or a fricative; otherwise, L tone.

[2]We conducted a supplementary linear mixed-effects analysis with stiffness (peak velocity normalized over displacement) as the dependent variable and focus condition as a fixed factor, as suggested by a reviewer. As this analysis revealed no substantial differences across focus types, we do not report these results in detail.

Aylett, M., and Turk, A. (**2004**). "The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech," Lang. Speech **47**(1), 31–56.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.

Baumann, S., Grice, M., and Steindamm, S. (**2006**). "Prosodic marking of focus domains—Categorical or gradient," in *Proceedings of the 3rd International Conference on Speech Prosody 2006*, pp. 301–304.

Beckman, M. E. (**1996**). "The parsing of prosody," Lang. Cogn. Proc. **11**(1-2), 17–68.

Beckman, M. E., and Edwards, J. (**1990**). "Lengthenings and shortenings and the nature of prosodic constituency," in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, edited by J. C. Kingston and M. E. Beckman (Cambridge University Press, Cambridge, UK), Vol. 152, p. 178.

Beckman, M. E., and Edwards, J. (**1994**). "Articulatory evidence for differentiating stress categories," in *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form* (Cambridge University Press, Cambridge, UK), pp. 7–33.

Beckman, M. E., Edwards, J., and Fletcher, J. (**1992**). "Prosodic structure and tempo in a sonority model of articulatory dynamics," in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, edited by G. J. Docherty and D. R. Ladd (Cambridge University Press, Camdrige, UK), pp. 68–89.

Beckman, M. E., and Pierrehumbert, J. B. (**1986**). "Intonational structure in Japanese and English," Phonol. Yearb. **3**, 255–309.

Breen, M., Fedorenko, E., Wagner, M., and Gibson, E. (**2010**). "Acoustic correlates of information structure," Lang. Cogn. Processes **25**(7-9), 1044–1098.

Byrd, D., Krivokapić, J., and Lee, S. (**2006**). "How far, how long: On the temporal scope of prosodic boundary effects," J. Acoust. Soc. Am. **120**(3), 1589–1599.

Byrd, D., and Saltzman, E. (**1998**). "Intragestural dynamics of multiple prosodic boundaries," J. Phon. **26**(2), 173–199.

Byrd, D., and Saltzman, E. (**2003**). "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening," J. Phon. **31**(2), 149–180.

Cho, T. (**2006**). "Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English," Lab. Phonol. **8**, 519–548.

Cho, T. (**2011**). "Laboratory phonology," in *The Continuum Companion to Phonology* (Continuum International Publishing Group, London/New York), pp. 343– 368.

Cho, T. (**2015**). "Language effects on timing at the segmental and suprasegmental levels," in *The Handbook of Speech Production* (Wiley-Blackwell, Hoboken, NJ), pp. 505–529.

Cho, T. (**2022**). "Linguistic functions of prosody and its phonetic encoding with special reference to Korean," in *Japanese/Korean Linguistics 29*, edited by K. Horie, K. Akita, Y. Kubota, D. Y. Oshima, and A. Utsugi [CLSI Publications (Stanford University), Stanford, CA], pp. 1–24.

Cho, T., and Keating, P. A. (**2001**). "Articulatory and acoustic studies on domain-initial strengthening in Korean," J. Phon. **29**(2), 155–190.

Cruttenden, A. (**1997**). *Intonation* (Cambridge University Press, Cambridge, UK).

de Jong, K. J. (**1991**). "The oral articulation of English stress accent," Ph.D. dissertation, The Ohio State University, Columbus, OH.

de Jong, K. J. (**1995**). "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation," J. Acoust. Soc. Am. **97**(1), 491–504.

Dimitrova, S., and Turk, A. (**2012**). "Patterns of accentual lengthening in English four-syllable words," J. Phon. **40**(3), 403–418.

Edwards, J., Beckman, M. E., and Fletcher, J. (**1991**). "The articulatory kinematics of final lengthening," J. Acoust. Soc. Am. **89**(1), 369–382.

Fletcher, J. (**2010**). "The prosody of speech: Timing and rhythm," in *The Handbook of Phonetic Sciences*, edited by F. Gibbon, W. J. Hardcastle, and J. Laver (Wiley, New York), pp. 521–602.

Fowler, C. A. (**1995**). "Acoustic and kinematic correlates of contrastive stress accent in spoken English," *Producing Speech: Contemporary Issues: For Katherine Safford Harris*, pp. 355–373.

Fox, A. (**2002**). *Prosodic Features and Prosodic Structure: The Phonology of 'supraseg- Mentals* (Oxford University Press, Oxford, UK).

Gracco, V. L. (**1988**). "Timing factors in the coordination of speech movements," J. Neurosci. **8**(12), 4628–4639.

Gussenhoven, C. (**1983**). "Testing the reality of focus domains," Lang. Speech **26**(1), 61–80.

Hawkins, S. (**1992**). "An introduction to task dynamics," in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, edited by G. J. Docherty and D. R. Ladd (Cambridge University Press, Camdrige, UK), pp. 9–25.

Hermes, A., Becker, J., Mücke, D., Baumann, S., and Grice, M. (**2008**). "Articulatory gestures and focus marking in German," In *Proceedings of the 4th International Conference on Speech Prosody 2008*, pp. 457–460.

Hyman, L. M. (**1978**). "Word demarcation," Univ. Hum. Lang. **2**, 443–470.

Iskarous, K., and Pouplier, M. (**2022**). "Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology," J. Phon. **95**, 101195.

Jang, J. (**2023**). "Functions of prosodic boundaries in an edge-prominence language: Marking focus and phrasing in Seoul korean," Ph.D. dissertation, University of California, Santa Barbara, CA.

Jang, J., and Katsika, A. (**2023**). "Edgy articulation: The kinematic profile of accentual phrase boundaries in Seoul Korean," in *Proceedings of Hanyang International Symposium on Phonetics and Cognitive Sciences of Language (HISPhonCog) 2023*, pp. 204–205.

Jang, J., and Katsika, A. (**2024**). "Phrase boundaries lacking word prosody: An articulatory investigation of Seoul Korean," J. Acoust. Soc. Am. **155**(5), 3521–3536.

Jeon, H.-S., and Nolan, F. (**2017**). "Prosodic marking of narrow focus in Seoul Korean," Lab. Phonol. **8**(1), 2–30.

Jun, S.-A. (**1993**). "The phonetics and phonology of Korean prosody," Ph.D. dissertation, Ohio State University, Columbus, OH.

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

Jun, S.-A. (**2000**). "K-ToBI (Korean ToBI) labelling conventions," Speech Sci. **7**(1), 143–170.

Jun, S. A. (**2005**). Prosodic Typology: The Phonology of Intonation and Phrasing (Oxford University Press, Oxford, UK)., Vol. 1.

Jun, S.-A. (**2011**). "Prosodic markings of complex NP focus, syntax, and the pre-/post-focus string," in *Proceedings of the 28th West Coast Conference on Formal Linguistics*, pp. 214–230.

Jun, S.-A. (**2014**). "Prosodic typology: By prominence type, word prosody, and macro-rhythm," in *Prosodic Typology II: The Phonology of Intonation and Phrasing* (Oxford University Press, Oxford, UK), pp. 520–539.

Jun, S.-A., and Lee, H.-J. (**1998**). "Phonetic and phonological markers of contrastive focus in Korean," in *Proceedings of the 5th International Conference on Spoken Language Processing*.

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., and Saltzman, E. (**2020**). "The role of focus in accentual lengthening in American English: Kinematic analyses," in *Proceedings of the 10th International Conference on Speech Prosody 2020*, pp. 275–279.

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., and Saltzman, E. (**2023**). "A hierarchy of prominence: The production and perception of focus in American English," in *Proceedings of the International Congress of Phonetic Sciences 2023*, pp. 1696–1700.

Katsika, A., and Tsai, K. (**2021**). "The supralaryngeal articulation of stress and accent in Greek," J. Phon. **88**, 101085.

Keating, P., and Shattuck-Hufnagel, S. (**2002**). "A prosodic view of word form encoding for speech production," UCLA Work. Papers Phon. **101**, 112–156.

Kim, J. J., Kim, S., and Cho, T. (**2024**). "Preboundary lengthening and articulatory strengthening in Korean as an edge-prominence language," Lab. Phonol. **1**(1), 9880.

Kopera, H. C., and Grigos, M. I. (**2025**). "How does prosodic prominence impact articulatory movement parameters and movement variability in adults who stutter?," J. Speech. Lang. Hear. Res. **68**, 3602–3625.

Ladd, D. R. (**1996**). *Intonational Phonology* (Cambridge University Press, Cambridge, UK).

Ladefoged, P. (**1967**). *Three Areas of Experimental Phonetics* (Oxford University Press, London,UK).

Ladefoged, P., and Loeb, G. (**2002**). "Preliminary studies on respiratory activity in speech," UCLA Work. Papers Phon. **101**, 50–60.

Lehiste, I. (**1970**). *Suprasegmentals* (MIT Press, Cambridge, MA).

Lenth, R. V. (**2023**). "emmeans: Estimated marginal means, aka least-squares means," R package, https://CRAN.R-project.org/package=emmeans (Last viewed June 23, 2025).

Liberman, M., and Prince, A. (**1977**). "On stress and linguistic rhythm," Ling. Inq. **8**(2), 249–336.

Mücke, D., and Grice, M. (**2014**). "The effect of focus marking on supralaryngeal articulation—Is it mediated by accentuation?," J. Phon. **44**, 47–61.

Munhall, K. G., Ostry, D. J., and Parush, A. (**1985**). "Characteristics of velocity profiles of speech movements," J. Exp. Psychol.: Human Percept. Perform. **11**(4), 457.

Nespor, M., and Vogel, I. (**1986**). *Prosodic Phonology* (Foris Publications, Bonn, Germany).

Ostry, D. J., and Munhall, K. G. (**1985**). "Control of rate and duration of speech movements," J. Acoust. Soc. Am. **77**(2), 640–648.

Pierrehumbert, J., and Beckman, M. E. (**1988**). *Japanese Tone Structure* (MIT Press, Cambridge, MA).

R Core Team (**2024**). "R: A language and environment for statistical computing," Vienna, Austria, https://www.R-project.org/ (Last viewed June 23, 2025).

Roessig, S., and Mücke, D. (**2019**). "Modeling dimensions of prosodic prominence," Front. Commun. **4**, 44.

Roessig, S., Winter, B., and Mücke, D. (**2022**). "Tracing the phonetic space of prosodic focus marking," Front. Artif. Intell. **5**, 842546.

Roon, K. D., Gafos, A. I., Hoole, P., and Zeroual, C. (**2007**). "Influence of articulator and manner on stiffness," in *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 409–412.

Saltzman, E., Nam, H., Krivokapic, J., and Goldstein, L. (**2008**). "A task-dynamic toolkit for modeling the effects of prosodic structure on articulation," In *Proceedings of the 4th International Conference on Speech Prosody 2008*, pp. 175–184.

Selkirk, E. (**1984**). *Phonology and Syntax: The Relationship between Sound and Structure* (MIT Press, Cambridge, MA).

Shattuck-Hufnagel, S., and Turk, A. E. (**1996**). "A prosody tutorial for investigators of auditory sentence processing," J. Psycholinguist. Res. **25**, 193–247.

Shin, S., Kim, S., and Cho, T. (**2015**). "What is special about prosodic strengthening in Korean: Evidence in lingual movement in v# v and v# cv," in *Proceedings of the 18th International Congress on Phonetic Science*.

Sorensen, T., and Gafos, A. (**2016**). "The gesture as an autonomous nonlinear dynamical system," Ecol. Psychol. **28**(4), 188–215.

van der Hulst, H. (**1999**). "Word accent," in *Word Prosodic Systems in the Languages of Europe* (Mouton de Gruyter, Berlin, Germany), pp. 3–116.

Venditti, J. J., Jun, S.-A., and Beckman, M. E. (**2014**). "Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English," in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, edited by J. L. Morgan and K. Demuth (Psychology Press, Hove, UK), pp. 299–324.