Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean

Taehong Cho & Sun-Ah Jun
UCLA

1. Introduction
Researchers have shown that realizations of segments are influenced by various prosodic factors such as stress (de Jong et al., 1993; Beckman & Edwards 1994; de Jong 1995) and phrase-final lengthening (Oller 1973, Klatt 1975, Wightman et al. 1992, Beckman & Edwards, 1994). For this reason, studying segmental properties requires one to control the prosodic conditions imposing on the segment. The other side of this observation is that prosodic information is cued by low-level phonetic details of segmental realization. For example, segments are lengthened at a phrase-final position, and speakers identify the end of a phrase boundary by means of the lengthening of phrase-final segments. Such dependency of segmental phonetics on prosodic structures has motivated quite a few studies to examine different kinds of prosodically determined locations which affect the acoustic realizations of a segment as well as the spatial and temporal properties of individual articulatory gestures (here by ‘gestures’ we mean simply measurable movements of articulators).

In addition to the effect of phrase-final position on segments, researchers have examined the effect of the phrase-initial position on segment realization. For example, Cooper (1991) found that English voiceless aspirated stops in word-initial position have a larger glottal opening gesture compared to word-medial position. Pierrehumbert & Talkin (1992) reports that the /h/ in English tomahawk at the beginning of an Intonational phrase has a longer VOT compared to phrasal medial position. Their acoustic studies also show, though indirectly, that English /h/ is produced with larger glottal gestures after an Intonational Phrase boundary compared to phrase-medial position. Along the same lines, Jun (1993, 1995) shows that VOT for Korean /h/ is longer phrase (Accentual Phrase) initially than medially. She also found that the lengthening was cumulative: phrase-initial VOT was longer than word initial VOT, which is longer than word medial VOT. These findings suggest that glottal gesture is strengthened domain-initially, resulting in a larger and longer glottal opening.

Such domain-initial effects have also been found in supralaryngeal articulations. Using electropalatography (EPG), a series of studies has shown that consonants are, in general, produced with greater and longer linguopalatal contact (contact between tongue and the palate area) domain-initially than domain-finally. For example, Fougeron and Keating (1997) found that the amount of linguopalatal contact for the alveolar nasal /n/ in reiterant English speech is greater when the consonant is domain-initial than when it is not. They also found the effect is cumulative — the higher the prosodic position, the more linguopalatal contact —, and termed this effect domain-initial strengthening. Similar findings have been
found in several other languages, including French (Fourgeron & Keating, 1996, Fougeron 1999), Taiwanese (Hsu & Jun, 1998; Hayashi, Hsu & Keating, 1999), Estonian (Gordon, 1996), and Korean (Cho, 1998; Cho & Keating 1999). (See also Keating, Cho, Fougeron & Hsu (1999) for a cross-linguistic study.)

However, it has not been clear exactly what is strengthened domain-initially. Among many possible explanations, the most commonly agreed-upon characteristic of the domain-initial effect is to enhance the contrast between the initial segment and its neighbors, i.e., syntagmatic contrast enhancement. Specifically, the domain-initial strengthening is viewed to be syntagmatically motivated, resulting in enhancement of CV (or VC) contrast. Existing data (e.g., Pierrehumbert & Talkin, 1992; Farnetani & Vayra, 1996; Fougeron & Keating, 1996, 1997; Hsu & Jun, 1998) seem to suggest that what is strengthened domain-initially is ‘consonantality’ of the segment, thus enhancing the syntagmatic contrast with the following vowel. For example, articulatorily, consonants become more consonant-like domain-initially by way of more extreme oral constriction. Acoustically, voiceless consonants become more consonantal by way of longer closure duration and longer VOT (Pierrehumbert & Talkin, 1992; Jun, 1993, 1995) which enhances voicelessness, thus being more contrastive with the following vowel. Fougeron (1998, 1999) found that nasal airflow of /n/ was smaller phrase-initially than phrase-medially. That is, by way of weaker nasal airflow, nasals become less sonorant, and thus more consonantal in domain-initial positions. Finally, Jun (1995) and Hsu & Jun (1996) found that the closure duration of a phrase-initial consonant is longer than a phrase-medial one, while the duration of the vowel following the phrase-initial consonant is shorter than that of phrase-medial vowel. This suggests that the scope of domain-initial lengthening is smaller than that of domain-final lengthening. The former appears to be limited to the initial segment of a phrase, while the latter is limited to the phrase-final syllable, sometimes extending even to the left of the final syllable (e.g., Hofhuis et al., 1995; Kohler, 1983). This also suggests that the domain-initial strengthening may be primarily motivated to emphasize the phrase-initial CV contrast.

On the other hand, there is another type of enhancement that consonants may undergo domain-initially. That is, domain-initial consonants can be strengthened by being more distinct from other types of consonants, i.e., paradigmatic contrast enhancement, a maximization of phonemic distinction of contrastive sounds. One such study that examines domain-initial strengthening systematically in terms of paradigmatic contrast enhancement is Hsu & Jun (1998). They examined variation of VOTs in Taiwanese for aspirated /kʰ/, unaspirated /t/, and voiced /b/, across different prosodic positions. As shown in Figure 1, they found that domain-initial /kʰ/ and /b/ have more positive and more negative VOTs, respectively, whereas the VOT for unaspirated /t/ does not change at all across prosodic positions. Based on these results, they suggested that some of features associated with domain-initial segments are enhanced in order to enhance paradigmatic contrast among segments (e.g., the aspiration for /kʰ/ and the voicing
for /b/), and that VOT for the unaspirated stop was not enhanced in any direction, because otherwise it would act negatively in the paradigmatic or phonemic voicing contrast among stops. They concluded that some features are enhanced paradigmatically, and some are syntagmatically, depending on the sound system of the language and the type of contrasts the language may choose to enhance.

![Figure 1. VOT duration (msec) for three stop categories at three prosodic positions. (From Hsu & Jun, 1998, p.82)](image_url)

The current study further explores the nature of domain-initial strengthening, based on aerodynamic data of voiceless stop triplets in Korean, and examines what type of enhancements (syntagmatic or paradigmatic or both) occurs domain-initially among these stops, especially in laryngeal articulation in comparison with oral articulation. In what follows, we will first provide a brief review of Korean stops necessary for further discussion in this paper, and report results of an aerodynamic study which investigates how the glottal gesture is strengthening domain-initially by examining the degree to which amount of oral airflow varies depending on prosodic positions and consonant types. Finally we will discuss domain-initial strengthening in terms of contrast maximization strategies.

1.1 Review of Korean stops

Korean has well-known three-way contrastive stops: strongly aspirated stop (ex. /tʰ/), moderately aspirated lenis stop (ex. /t/) and fortis stop (ex. /t/). One of the most important cues that differentiate these stops from one another is VOT (e.g., Lisker & Abramson 1964; Han & Weitzman, 1970; Silva 1992; M-R Kim 1994, Han 1996; Cho, 1996). In general, VOT is shortest for the fortis stop, intermediate for the lenis stop, and longest for the aspirated stop. It has been also reported that the glottal opening patterns similarly with VOT, being smallest for the fortis stop,
intermediate for the lenis stop, and largest for the aspirated stop (e.g., Kagaya, 1974; Jun, Beckman & Lee, 1998).

As we briefly introduced earlier, VOT varies as a function of prosodic position (e.g., Jun, 1993, 1995; Cho & Keating, 1999). Figure 2 shows variation in VOT for the aspirated stop /pʰ/ across three prosodic positions based on Jun (1993): VOT is longer phrase-initially, shorter word-initially, and even shorter word-medially. The longer VOT for the phrase-initial aspirated stop can be interpreted as being due to a greater glottal opening domain-initially. Jun, Beckman & Lee (1998) observed variations in the glottal opening area using a fiberscopic technique, and confirmed that for the voiceless aspirated stop, the glottal area is longer phrase-initially than phrase-medially. The larger and longer glottal opening domain-initially then can be viewed as an instance of syntagmatic CV enhancement in that the resulting greater aspiration enhances voicelessness of the stop, being further contrastive with the following vowel which is voiced.

![Figure 2. VOT of Korean /pʰ/ as a function of prosodic position. (Based on Jun 1993, p.235).](image)

Now, moving on to the supralaryngeal level, evidence of domain-initial strengthening in oral articulation of three-way contrastive stops is found in Cho (1998) and Cho & Keating (1999). They examined linguopalatal contact for all three stops (/t, tʰ, tʰ/) and nasal /n/, and found that each consonant has larger and longer linguopalatal contact domain-initially than domain-medially, as is evident in Figure 3. (Note that the Utterance-initial (Ui) is not different from the Intonational Phrase-initial (IPi) in terms of intonational patterns. Ui and IPi differ in that Ui is sentence-initial, preceded by a substantial pause (marked by ‘.’ in orthography), while IPi is sentential-medial, preceded by a medium length pause and a boundary tone (marked by ‘,’ in orthography.)

Note that linguopalatal contact (or oral constriction) is greater domain-initially than domain-medially in a cumulative fashion, regardless of consonant type. Such greater linguopalatal contact in domain-initial positions can be considered as syntagmatic enhancement of CV contrast. Cho & Keating also found
that consonants can be generally differentiated by linguopalatal contact, giving a pattern of \( t^* > t^b > t \), especially in lower prosodic positions. What is interesting is that such contrast among these three stops becomes less distinctive as the prosodic position moves up in the hierarchy, and eventually becomes blurred in the Utterance-initial position, as can be seen in Figure 3. This clearly suggests that paradigmatic contrast in oral constriction among coronal consonants is not enhanced at all domain-initially, but rather it becomes fuzzy at the expense of the apparent syntagmatic CV enhancement.

Figure 3. Variation in linguopalatal contact as a function of prosodic position. Error bar refers to 97% confidence intervals. Ui=Utterance-initial, IPi=Intonational Phrase-initial, API=Acccentual Phrase-initial, Wi=Word-initial. (Redrawn based on Figure 6 in Cho & Keating 1999.)

1.2 Predictions

If domain-initial articulation is limited to syntagmatic contrast enhancement as we have observed in VOT (for the aspirated stop only) and linguopalatal contact in the preceding section, it can be further hypothesized that glottal gesture will be strengthened domain-initially in one direction regardless of consonant types, resulting in a greater opening of the glottis for all stops, as schematized in Figure 4a. Alternatively, if domain-initial strengthening plays a primary role in enhancing paradigmatic contrast, we would expect that different stops have different patterns in glottis opening in a way that optimizes paradigmatic contrast among them, as schematized in Figure 4b.
2. Method

In order to test the two competing hypotheses, aerodynamic data were collected from four male and one female Seoul Korean speakers as outlined below.

The test consonants are three bilabial stops /p, pʰ, p*/ (where /p*/ represents the fortis, or tense, stop, for which there is no official IPA transcription). Bilabials are used because oral pressure behind lips is the easiest to measure. Each test consonant was placed in a fixed segment context within a set of sentences. The corpus is given in (1):

(1) a. Intonational Phrase-initial (IPi)
penin ikəjo  The ship is this one
pʰenin ikəjo  The card is this one
p*enin ikəjo  The bone is this one

b. Accentual Phrase-initial (API)
ikəsin pejo  This is a ship
ikəsin pʰejo  This is a card
ikəsin p*ejo  This is a bone

c. Word-initial (Accentual Phrase-medial)
ikəsin nepəjo  This is my ship
ikəsin nepʰejo  This is my card
ikəsin nep*ejo  This is my bone

The sets of sentences were constructed to vary in their likely phrasing, so that the prosodic context of the test consonants would vary. When subjects produced these sentences with the expected phrasings, then the test consonant was initial in a prosodic domain that varies from IP to W. In the present study,
following Jun (1993, 1998, 2000) and Beckman and Jun (1996) we adopt the intonationally defined prosodic hierarchy as shown in Figure 5. In this model, Syllables (S) are grouped into Words (W); Words are grouped into Accentual Phrases (AP); Accentual Phrases are grouped into Intonational Phrases (IP). Following the Strict Layer Hypothesis (Selkirk 1986), it is assumed that the beginning and end of each higher domain is also the beginning and end of lower domains. (Note that in our study the IP-initial is also the initial position of the sentence, which is equivalent to the Utterance-initial in Cho & Keating (1999).)

![Prosodic Structure of Korean (adopted from Jun 1993, 2000 and Beckman and Jun 1996).](image)

Each sentence in (1) was repeated nine times. In order to obtain reasonably consistent prosody for each sentence type without over instruction, sentences were not randomized. Instead, for a given test consonant, a subject produced 3 or 4 repetitions of one sentence in a block. This gave a total of 405 sentences to examine (3 stops by 3 prosodic positions by 5 speakers by 9 repetitions).

Oral airflow and pressure were recorded using the Macquiler X16 system (Scicon). Speakers held a face mask against the lower part of the face, below the nose, capturing all the oral airflow. They also held a tube (internal diameter 2 mm) between their lips to record the pressure of the air in the mouth. A microphone embedded in the face mask recorded the audio signal. The flow and pressure signals were sampled at a rate of 2 kHz and the audio signal was sampled at 10 kHz.

The maximum airflow after the release of the closure and the peak oral pressure during the closure were measured, as indicated by the arrows (a) and (b) in Figure 6, respectively. In addition, integrated airflow (i.e., the area below airflow contour) was calculated over the period of time from the release of the closure to the onset of the vowel. Finally, VOT was measured from the release of the closure to the onset of the vowel, the same period of time used to calculate the integrated airflow. These measurements were used as indicative of how large and long the glottal opening would be.
3. Results

3.1 VOT

Results of a repeated measures ANOVA ([prosodic position] by [consonant type]) showed that there are main effects of [prosodic position] (F(2, 8) = 150.974, p < .0001) and [consonant type] (F(2, 8) = 1306.453, p < .0001). Fisher's PLSD posthoc pairwise comparison confirmed that all domains are differentiated from one another by VOT (at a significance level of .05) in increasing order of Wi, API, and IPi. In addition, VOT varies in increasing order of $p^* < p < p^h$ at a significance level of $p < .0001$. However, there is a significant [prosodic position by consonant type] interaction (F(4, 16) = 51.778, p < .0001). As is evident in Figure 7, the interaction is mainly due to the fact that VOTs for the fortis stop (/p*/) do not vary substantially across prosodic positions, while VOTs for both aspirated and lenis stops (/p^h, p/) increase to a great degree as a position moves up in the prosodic hierarchy. In fact, mean VOT for /p*/ is slightly shorter in a higher position than in a lower position.

Figure 7. Variation in VOT as a function of prosodic positions.
3.2. Airflow

There is a main effect of [prosodic position] in both peak airflow and integrated airflow (for peak airflow $F(2, 8) = 49.767, p < .0001$; for integrated airflow $F(2, 8) = 56.199, p < .0001$). There is also a main effect of [consonant type] in both parameters (for peak airflow $F(2, 8) = 591.365, p < .0001$; for integrated airflow $F(2, 8) = 467.284, p < .0001$). However, significant interactions were found between [prosodic position] and [consonant type] (for peak airflow, $F(4, 16) = 24.926, p < .0001$; for integrated airflow, $F(4, 16) = 18.009, p < .0001$).

Figure 8 shows variations as a function of prosodic positions in (a) peak airflow and (b) integrated airflow. As was the case with VOT, the interactions are mainly because both the peak airflow and integrated airflow for the fortis stop are nearly invariable, whereas those for the other two types of stops vary with prosodic positions. Posthoc comparisons made separately for each stop showed a pattern of $IPi > APi > Wi$ at a significance level of $p < .0001$ in the integrated airflow; but in the peak airflow, such a three-way distinction was made only for the lenis $/p/$. For the aspirated $/p^*/$, significant difference was found only between $IPi$ and $Wi$.

![Figure 8. Variation in (a) peak airflow and (b) integrated airflow as a function of prosodic positions.](image)

While all speakers showed a similar pattern of variations as a function of prosodic positions for $/p^h/$ and $/p/$, a somewhat inconsistent pattern was found for $/p^*/$ across speakers. However, the range of variation in integrated airflow for the fortis stop (about 6 liter/sec averaged across speakers) seems too small to indicate any systematic direction, as compared with those for the lenis and aspirated stops (about 167 and 431 liter/sec, respectively). Nonetheless, the mean peak airflow and integrated airflow are slightly smaller $IP$-initially than $Wi$, which at least guarantees that glottal opening is not larger $IP$-initially than $domain$-medially. (In fact, three out of five speakers showed consistently lower integrated airflow $IP$-initially than $Word$-initially, though statistically not significant.)
4. Discussion

Thus far we have examined variations in VOT and airflow as a function of prosodic positions. Overall, results reveal a similar pattern for VOT and airflow, showing that values in both parameters are higher in a higher prosodic position (IPi) for the lenis and aspirated stops, while there is a tendency towards lowered VOT and airflow for the fortis stop in a higher prosodic position.

One of the key questions raised in this study was whether laryngeal articulation is strengthened syntagmatically or paradigmatically among three-way contrastive stops. Overall, results indicate that laryngeal articulation is strengthened, but not in a uniform way to enhance only syntagmatic contrast among neighboring segments. While the greater VOT and airflow in a higher prosodic position suggest that the glottis is larger domain-initially than domain-medially, the inverse tendency for the fortis stop suggests that the glottis is smaller, being more constricted in a higher prosodic position. (Given that there is some inconsistency across speakers, we still can assure at least no larger glottal opening for the fortis stop domain-initially.) In fact, the observed inverse tendency in aerodynamic data can be further supported by evidence in Jun, Beckman & Lee (1998) who made direct observations of the change in the glottal area across prosodic positions, using a fibrescopic technique. Our closer examination of their data reveals that the overall peak glottis area during the closure and the glottis opening at the time of the release tend to be smaller AP-initially than AP-medially. Figure 9 shows an example pair in which the word /k*it[i]/ (‘the end-nominative’) occurs AP-initially and AP-medially.

![Figure 9](image)

Figure 9. Areas of glottal opening for the fortis stop /k*/ in /k*it[i]/. The horizontal axes are time units (each tick equals 16.6ms) and the vertical axes are the glottal opening area on an arbitrary scale. ‘R’ refers to the release of closure, and ‘V,’ to the voicing onset of the following vowel. Excerpted from Figure 10, p. 63 in Jun et al. (1998).
The asymmetric domain-initial effects on laryngeal articulation among different stops can be accounted for by viewing the domain-initial strengthening as enhancement of laryngeal features. The three-way contrastive stops (lenis /p/, fortis /pʰ/, aspirated /pʰ/) in Korean can be phonologically differentiated by two privative laryngeal features ([spread glottis] and [constricted glottis]) as shown in (2):

(2) (following Lombardi, 1991a, b)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirated</td>
<td>[spread glottis]</td>
</tr>
<tr>
<td>Fortis</td>
<td>[constricted glottis]</td>
</tr>
<tr>
<td>Lenis</td>
<td>unspecified</td>
</tr>
</tbody>
</table>

On the one hand, the increased VOT and airflow for the domain-initial aspirated stop can be interpreted as enhancement of the feature [spread glottis]. On the other hand, the reversed tendency for the domain-initial fortis stop may be seen as enhancement of the feature [constricted glottis]. Such enhancement of laryngeal features in a domain-initial position appear to maximize paradigmatic contrast among stops. As can be seen in Figures 7 and 8, it is the highest domain-initial position (IPii) where stops are maximally dispersed along the VOT and integrated airflow. These patterns are consistent with the prediction schematized in Figure 4b.

Finally, the lenis stop, being unspecified for both features still shows an increased VOT and airflow domain-initially. This indicates that, when a stop has no specification for a distinctive feature, the stop is nonetheless strengthened, but this time presumably to enhance the consonantality that results in a greater syntagmatic CV contrast. Interestingly, the range of variation for the lenis stop rarely overlapped with that of the aspirated stop, implying that such syntagmatic enhancement may be limited to an extent which does not blur the paradigmatic contrast—e.g., between the aspirated and lenis stops.

5. Closing remarks

Findings in the present study together with existing data appear to suggest that domain-initial strengthening can be viewed as a complex linguistic phenomenon. It engenders not only syntagmatic contrast enhancement between neighboring segments (e.g., CV contrast enhancement), but also paradigmatic contrast maximization among phonemically contrastive sounds (e.g., fortis vs. lenis vs. aspirated stops). We believe that domain-initial strengthening provides perceptual cues for the prosodic structure and information groupings by maximizing contrast, syntagmatic or paradigmatic, of domain-initial segment in multiple acoustic and articulatory dimensions. That is, domain-initial strengthening seems to help listeners to segment the incoming flow of the speech into smaller units, and recover the meaning of the utterance and the speaker's intention. Such an enhancement strategy may also facilitate lexical access to the domain-initial lexical item which generally has less contextual or discourse information than lexical items occurring later domain-medially.
However, given only a handful of data available, we cannot be assured that such effects are universally manifest across languages. As noted in Hsu & Jun (1998), it is conceivable that some features are enhanced paradigmatically and some are syntagmatically, depending on the sound system of the language and the type of contrast the language may choose to enhance. Much work remains to be done cross-linguistically before making generalizations about the linguistic role that domain-initial strengthening plays and how strengthening is realized.

Acknowledgments

We would like to thank our subjects for participating in the experiments. Special thanks go to Peter Ladefoged, who helped us with the aerodynamic equipment settings, data acquisition and data analysis; and to Henry Teheranizeh who made the measurement of the integrated airflow possible in the Mcquirer system. We also thank Leston Buell for proofreading and valuable suggestions; and the audience of the CLS 36 and other members of the UCLA Phonetics Lab for comments on earlier presentations of the work.

Notes

1 The word [p*ere] is not a standard Seoul Korean word, but Korean speakers have no trouble pronouncing it when written in Hangul orthography)

2 There are several alternative accounts. For example, C. Kim (1965) and Kim-Renaud (1974) proposed two features [+/- tensity] and [+/- aspiration]; Halle & Stevens (1971) use four features, [+/- spread glottis], [+/- constricted glottis], [stiff vocal folds] and [slack vocal folds]; K-H. Kim (1987) uses [+/- spread glottis] and [+/- constricted glottis] in the framework of underspecification theory. See also Silva (1992) for a different approach using Steriade's (1993) Aperture Theory.

References


Beckman, Mary and Jun, Sun-Ah. 1996 K-ToBI (Korean ToBI) Labeling Conventions. Unpublished manuscript. Ohio State University & UCLA.


