



How does focus-induced prominence modulate phonetic realizations for Korean word-medial stops?

Jiyoun Choi*

Department of Social Psychology, Sookmyung Women's University, Seoul, Korea

Abstract

Previous research has indicated that the patterns of phonetic modulations induced by prominence are not consistent across languages but are conditioned by sound systems specific to a given language. Most studies examining the prominence effects in Korean have been restricted to segments in word-initial and phrase-initial positions. The present study, thus, set out to explore the prominence effects for Korean stop consonants in word-medial intervocalic positions. A total of 16 speakers of Seoul Korean (8 males, 8 females) produced word-medial intervocalic lenis and aspirated stops with and without prominence. The prominence was induced by contrast focus on the phonation-type contrast, that is, lenis vs. aspirated stops. Our results showed that F0 of vowels following both lenis and aspirated stops became higher when the target stops received focus than when they did not, whereas voice onset time (VOT) and voicing during stop closure for both lenis and aspirated stops did not differ between the focus and no-focus conditions. The findings add to our understanding of diverse patterns of prominence-induced strengthening on the acoustic realizations of segments.

Keywords: intervocalic stops, lenis stops, aspirated stops, prominence, focus

1. Introduction

Existing research recognizes the critical role played by prosodic structure in understanding sound patterns in languages (Cho & McQueen 2005; de Jong 1995, 2004; de Jong & Zawaydeh, 2002; Fougeron & Keating, 1997; Turk & Shattuck-Hufnagel, 2007). The prosodic structure has been concerned with those elements of speech that mark boundaries of prosodic constituents (e.g., Intonational Phrase, Accentual Phrase) and prominence in the utterances (e.g., driven by new information as compared to given information). Depending on where in the prosodic structure phonemes occur, the

same phonemes would be phonetically differently produced, such that those segments that occur in phrase-initial positions (vs. phrase-medial positions) or those occurring with prominence (vs. without prominence) would be realized with spatial and/or temporal expansions.

It has been further documented that the prominence-induced strengthening is not uniform across languages but conditioned by language-specific sound system (e.g., Cho & McQueen, 2005; de Jong, 1995, 2004; de Jong & Zawaydeh, 2002; Kim et al., 2018). For instance, it has been shown that voice onset times (VOTs) are not uniformly lengthened in prosodically strong positions but rather

* jiyoun.choi@sookmyung.ac.kr, Corresponding author

Received 15 November 2020; Revised 14 December 2020; Accepted 14 December 2020

© Copyright 2020 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

they are modulated in language-specific ways, such that voiceless stops in Dutch are produced with shortened VOTs when the stops were under prominence while English voiceless stops showed lengthened VOTs under prominence (Cho & McQueen, 2005; Kim et al., 2018).

In the present study, we explore how Korean lenis and aspirated stops in word-medial positions would be realized as a function of prominence, to increase our understanding of prominence-induced strengthening by including the new target stops. To date, research on the prominence effects on Korean stops has tended to focus on stops in word- and phrase-initial positions. Kang & Guion (2008), for instance, showed that young speakers increased F0 for the word- and phrase-initial lenis and aspirated stops in clear speech (i.e., foreigner-directed speech) as compared to those in conversational speech. Choi et al. (2020) also showed higher F0, as well as longer VOT, for the word- and phrase-initial lenis and aspirated stops when the stops were focused than when they were not. Note that in both studies, the aspirated and lenis stops differed in the degree to which their F0 and VOT values were manipulated by clear or focused speech, and as a result, for instance, the F0 distance between the stops became larger in clear and focused speech than in conversational and unfocused speech, respectively.

There has been, however, very little research directly investigating prominence effects on intervocalic word-medial lenis and aspirated stops, although acoustic realizations of stop consonants substantially differ between in the intervocalic word-medial positions and in the word- and phrase-initial positions. In terms of VOT, for instance, the word-medial intervocalic lenis stops are likely to be produced as voiced, a phenomenon known as Intervocalic Lenis Stop Voicing (Jun, 1993), while the lenis stops in the word- and phrase-initial positions are voiceless. Regarding F0, high tones vs. low tones are categorically assigned to the word- and phrase-initial aspirated vs. lenis stops respectively by the intonational phonology (Jun, 1993, 1998, 2000), while in the word-medial (and thus not phrase-initial) positions there are no such tonal specifications that the intonational phonology categorically assigns as a function of whether it is an aspirated or lenis stop.

Given the substantial differences in F0 and VOT realizations of stops depending on the positions, in the present study, we explore how acoustic realizations of the word-medial intervocalic lenis and aspirated stops would be modulated by prominence. Along with the manipulations of the prominence, we compare male and female speakers to see whether and how the phonetic realizations for the stops may differ as a function of speakers' gender. A number of studies have demonstrated gender differences in VOT and F0 for lenis versus aspirated stops in word- and phrase-initial positions, with larger VOT and smaller F0 distances between the stops for males than females (Bang et al., 2018; Choi et al., 2020; Kang, 2014). Thus, to see whether there would be any gender differences for the word-medial intervocalic stops as well, we collect data from both female and male speakers.

2. Methods

2.1. Participants and recording

A total of 16 native speakers of Seoul Korean participated in the recording (age 19–24 years, mean=22 years at the time of testing, 8 females). All had been born and raised in Seoul or Gyeonggi

province (the area surrounding Seoul). The participants were not aware of the purpose of the study and were paid for the participation. The acoustic data were recorded at a sampling rate of 44 kHz, using a Tascam Hd-P2 digital recorder and a SHURE KSN44 microphone, in a sound attenuated booth at Hanyang Institute for Phonetics and Cognitive Science of Language, Seoul.

2.2. Speech materials and procedure

The tested contrast was a Korean aspirated vs. lenis stop contrast in alveolar place of articulations (*t^h/vs./t*). The following vowel was always low vowel /a/. A total of eight minimal pairs of disyllabic words were prepared, with the contrastive syllables occurring in word-internal positions (see Table 1 for all minimal pairs).

Table 1. Minimal pairs and English translations

Aspirated	Lenis
[pit ^h an] grief	[pitan] silk
[sat ^h an] Satan	[satan] division
[sat ^h an] candy	[satan] shrine
[sut ^h a] hand-making noodles	[suta] chatter
[sut ^h al] exploitation	[suta] otter
[tʃit ^h an] criticism	[tʃitan] egg garnish
[hʌt ^h an] vain effort	[hʌtan] sloppy
[jʌnt ^h an] briquette	[jʌntan] podium

The test words were then included in sentences such that the focus types (focus vs. no focus) could be manipulated. Example sets of sentences bearing a test word /pit^han/ (*grief*) are given in Table 2. Each set consisted of question and answer sentences.

To induce the intended prosodic structure as naturally as possible, we created a board game situation in which a participant conversed about where to put word cards. In this game, participants were first shown a correct answer picture on a computer screen. After that, they were presented with a pre-recorded question of where to put the word card, and then answered the question according to the answer picture (Figure 1). The questions were recorded by four speakers. The participants heard one of the four versions of each question, with an equal number of each version.

Table 2. Example sentences with the test word /pit^han/ as a function of focus type. The target word is underlined and focused items are in bold

Focus	Q ipʌne p*opin tanʌnin, [ip pitan ape nonayo]? “Shall I put this card in front of pitan ?”
	A aniyo. ipʌnen, [ip pit^han ape noayo]. “No, put it in front of pit^han this time.”
No focus	Q ipʌne p*opin tanʌnin, [ip pit^han ape nonayo]? “Shall I put this card in front of pit ^h an?”
	A aniyo. ipʌnen, [ip pit^han tye noayo]. “No, put it in back of pit ^h an this time.”

For the focus condition, for instance, the participants were given the answer picture as in Figure 2. The picture showed that the spot where the word card should be placed is not in front of /pitan/ (*silk*) as indicated by a red cross but in front of /pit^han/ (*grief*) as indicated by a green circle. Then the participants were presented with the pre-recorded question, “Shall I place this card in front of /pitan/ (*silk*)?” in which the wrong place was uttered with focus. The participant was then asked to answer the question by correcting the

wrong information as in “No, put it in front of /pit^han/(grief) this time.” As the correct answer (/pitan/ (silk)) differed from the word to be corrected (/pitan/ (silk)) only in terms of the phonation types of the target stops (/t/ vs. /t^h/), a contrast focus on the phonation-type contrast was induced. For the no-focus condition, the locus of information to be corrected was not the test word but direction (front vs. back) (see Figure 2), so that the test word was not focused. With respect to the phrase boundary, an IP-initial boundary was induced by having the adverb /ipʌnən/ (this time) just before the test word. The participants in fact placed the IP boundary between the adverb and the test word without difficulty.

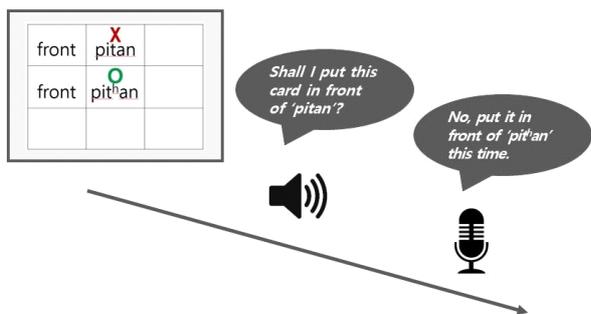


Figure 1. An illustration of a test trial

Prior to actual recording, the practice session was carried out. Only during the practice, the question and answer sentences were presented along with the answer pictures; during the actual recording, the sentences were not provided to help induce more spontaneous speech by preventing reading the written sentences. The entire test set was repeated two times in a randomized order. A total of 1,024 tokens (16 speakers×16 target words×2 focus types×2 repetitions) were obtained.

Focus			No focus		
front	X pitan				
front	O pit ^h an		X front	pit ^h an	O back

Figure 2. Example answer pictures with the test word /pit^han/ as a function of focus type

2.3. Measurements

F0 was measured at the midpoint (50%) of vowels following the target stops. We first used a Praat script (range of 65–500 Hz; time step of 10 ms) and then manually checked based on visual inspection of the pitch contour for each token. The F0 measurements in Hz were converted into semitones (St) using the formula $12[\log_2(\text{Hz}/100)]$ with a reference F0 of 100 Hz.

VOTs of target stops were measured from the stop release to the

voice onset of the following vowel defined as the onset of first formant (F1) seen in spectrograms. To examine whether and to what extent focus may affect degree of voicing during the stop closure (especially for lenis stops given the aforementioned lenis voicing rule), we calculated the percentage of the voiced interval (i.e., negative VOT) during the closure duration (%-Voicing, henceforth).

3. Results

A total of 29 tokens (2.8% of data) that were realized with inadequate prosodic patterns or that were difficult to measure acoustic parameters were discarded for analysis. Analysis was carried out in R (R Development Core Team, 2008) using linear mixed-effects models for each acoustic measure (F0, %-Voicing, VOT). Using sum contrast, Stop (0.5: aspirated, -0.5: lenis), Focus (0.5: focus, -0.5: no focus) and Gender (0.5: female, -0.5: male) were coded¹. The factors were included as fixed effects along with all interactions. The models included the maximal random effects structure supported by the data. Table 3 provides all fixed effect coefficients along with model formulae.

3.1. F0

There were significant main effects of Stop ($\beta=1.1$, $p<.05$) and Focus ($\beta=5.7$, $p<.001$), indicating that F0 was higher for aspirated than lenis stops and higher for focus than no-focus conditions (Figure 3). Not surprisingly, F0 was significantly higher for female than male speakers as well ($\beta=11.4$, $p<.001$). No interaction effects were observed.

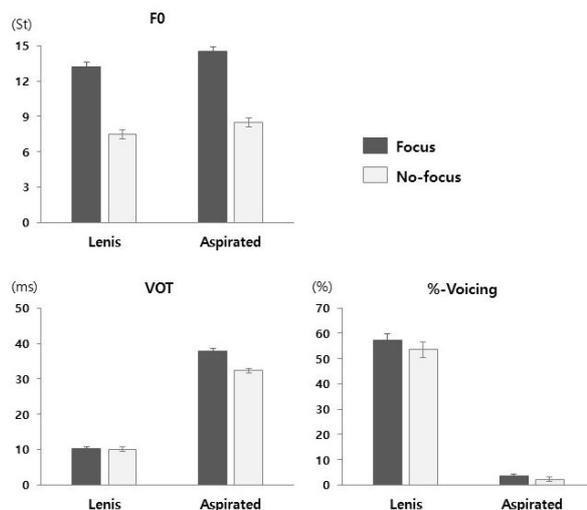


Figure 3. Means of F0, VOT, and %-Voicing for lenis and aspirated stops as function of focus condition. Error bars show standard errors.

¹ Instead of dummy coding which provides simple effects, sum coding was used to obtain main effects and interactions.

3.2. VOT

There was a significant main effect of Stop ($\beta=24.9, p<.001$), showing that VOT was longer for aspirated stops than lenis stops (Figure 3). Although VOT was numerically lengthened under focus especially for aspirated stops (5.4 ms; lenis stops = 0.1 ms), there was neither a main effect of Focus nor interactions including Stop and Focus. Other effects were also non-significant.

3.3. %-Voicing

We first checked how many tokens of lenis and aspirated stops were realized as voiced. Following the suggestion by Abramson & Whalen (2017), we counted tokens as voiced if there was voicing in more than 50% of the closure. As shown in Table 4, overall, over half of lenis stops were realized as voiced, and the proportion of being voiced for the lenis stops was higher for male than female speakers (83% vs. 41%). For the aspirated stops, almost none of the tokens were counted as voiced across all conditions.

Table 4. Numbers (percent proportion in parentheses) of stops that were realized as voiced or voiceless

		Lenis		Aspirated	
		Voiced	Voiceless	Voiced	Voiceless
Focus	M	88 (70%)	37 (30%)	2 (2%)	120 (98%)
	F	46 (37%)	80 (63%)	1 (1%)	127 (99%)
No focus	M	78 (62%)	47 (38%)	2 (2%)	123 (98%)
	F	54 (45%)	67 (55%)	2 (2%)	121 (98%)
Overall		266 (54%)	231 (46%)	7 (1%)	491 (99%)

Turning to results from the mixed-effects model on %-Voicing, a significant main effect of Stop was observed ($\beta=-52.2, p<.001$), indicating greater %-Voicing for lenis than aspirate stops (Figure 3). A main effect of Gender ($\beta=-11.1, p<.05$) and an interaction between Gender and Stop ($\beta=21.0, p<.001$) were also significant, showing that overall, %-Voicing was higher for the male speakers than female speakers but the %-Voicing difference between male and female speakers was reduced for the aspirated stops (0.6%, $-0.6=-11.1+21.0\times(0.5)$) as compared to the lenis stops (21.6%, $-21.6=-$

$-11.1+21.0\times(-0.5)$) by 21%. All the effects involving Focus were not significant.

4. Discussion and Conclusion

In the present study, we explored how word-internal intervocalic aspirated and lenis stops in Seoul Korean would be differently realized as a function of focused vs. unfocused conditions, together with whether and how the phonetic realizations of the stop contrasts would vary by speakers' gender. Several findings emerged from the study, which are summarized and discussed as follows.

First, our results showed that F0 of the vowels following the aspirated and lenis stops became higher under focus than no-focus conditions. The focus effect of increasing F0 is consistent with Yun (2013) which found higher F0 for intervocalic lenis stops when phrases with the lenis stops in the phrase-medial positions were focused than when not focused. It has been also reported that F0 for both aspirated and lenis stops in word- and phrase-initial positions became higher with focus than without focus (Choi et al., 2020), demonstrating the similar focus effects on F0 between for the word-medial intervocalic stops and for the word- and phrase-initial stops.

Unlike F0, we did not find any robust focus effects on VOT and %-Voicing for both aspirated and lenis stops. The finding again accords with that of Yun (2013) who also found no focus effect on VOT for the intervocalic lenis stops. Studies on word- and phrase-initial stops, however, demonstrated lengthened VOT under focus. For instance, Cho et al. (2011) examined word-initial aspirated stops as a function of focus and showed that VOT of the stops became longer when the stops were focused than when they were not. In Kang & Guion (2008), the effects of speaking style (clear vs. conversational speech) were exploited for word- and phrase-initial aspirated and lenis stops, and their results also showed longer VOT in clear than conversational speech for the aspirated stops. That is, our finding with no clear focus effects on VOT differs from earlier studies on word- and phrase-initial stops, showing dissimilar focus effects on stops depending of their positions.

Finally, our results showed that male speakers produced the intervocalic lenis stops as voiced more often than female speakers, indicating a greater degree of lenition for males than females. This finding is consistent with that of Sohn & Ahn (2011) who found the greater percentage of reduction in VOT from word-initial to

Table 3. Results from the linear mixed-effect models and final model formulae

	F0			VOT			%Voicing		
	Estimate	SE	t-value	Estimate	SE	t-value	Estimate	SE	t-value
(intercept)	10.9	0.5	22.3***	22.6	1.3	17.3***	28.9	3.1	9.1***
Stop	1.1	0.5	2.1*	24.9	2.7	9.3***	-52.2	5.1	-10.0***
Focus	5.7	0.8	7.2***	2.8	1.5	1.9	2.9	5.3	0.55
Gender	11.4	0.8	13.6***	-2.9	2.2	-1.3	-11.1	4.0	-2.7*
Stop×Focus	0.2	1.0	0.2	5.7	3.0	1.9	-2.6	10.3	-0.2
Stop×Gender	-0.2	0.2	-0.7	-9.3	4.5	-2.1	21.0	3.7	5.5***
Gender×Focus	0.0	1.2	0.0	-2.1	1.1	-1.9	-6.1	4.5	-1.3
Stop×Focus×Gender	-0.3	0.3	-1.1	-3.7	2.2	-1.7	8.9	7.5	1.1

Note. See below for the final model formulae.

F0~Stop×Focus×Gender+(1+Stop+Focus|speaker)+(1|item).

VOT~Stop×Focus×Gender+(1+Stop|speaker)+(1|item).

% Voicing~Stop×Focus×Gender+(1+Focus|speaker)+(1+Gender|item).

* $p<0.05$, *** $p<0.001$.

word-medial lenis stops for males than for females. They interpreted the finding as evidence demonstrating that male speakers are more prone to the weakening process in the intervocalic positions than females.

To conclude, this study has shown the focus effects on word-medial intervocalic aspirated and lenis stops, which differ from those of the stops in word- and phrase-initial positions. The findings add to the growing body of research that shows the varied patterns of prominence-induced strengthening, and suggest that it may be important to take into account the environments where phonemes occur to fully understand prominence-related phonetic modulation.

References

- Abramson, A. S., & Whalen, D. H. (2017). Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, 75-86.
- Bang, H. Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T. J. (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, 66, 120-144.
- Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, 39, 344-361.
- Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33, 121-157.
- Choi, J., Kim, S., & Cho, T. (2020). An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PLOS ONE*, 15(10), e0240682.
- de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97(1), 491-504.
- de Jong, K. J. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, 32, 493-516.
- de Jong, K. J., & Zawaydeh, B. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30, 53-75.
- Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101(6), 3728-3740.
- Jun, S. A. (1993). *The Phonetics and Phonology of Korean Prosody* (Doctoral 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) labeling conventions: version 3. *Speech Science*, 7(1), 143-169.
- Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of Acoustical Society of America*, 124, 3909-3917.
- Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 77-90.
- Kim, S., Kim, J., & Cho, T. (2018). Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics*, 71, 65-80.
- R Development Core Team. (2008). R: A language and environment for statistical computing [Computer program]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Sohn, H., & Ahn, M. (2011). Voice onset time differences in Daegu Korean stops as a function of their position in words. *The Journal of Linguistic Science*, 59, 149-180.
- Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35(4), 445-472.
- Yun, H. (2013). Segmental acoustic correlates associated with the Korean lenis stops. *Language Research*, 49, 73-94.

• **Jiyoun Choi**, Corresponding author
 Assistant Professor, Department of Social Psychology,
 Sookmyung Women's University,
 Cheongpa-ro 47-gil 100 Yongsan-gu, Seoul, 04310, Korea
 Tel: +82-2-2077-7331
 Email: jiyoun.choi@sookmyung.ac.kr
 Fields of interest: psycholinguistics, phonetics