Voice quality distinctions of the three-way stop contrast under prosodic strengthening in Korean[*]

Jiyoung Jang[1] · Sahyang Kim[2] · Taehong Cho[1,3,**]

[1]Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, Korea
[2]Department of English Education, Hongik University, Seoul, Korea
[3]Department of English Language and Literature, Hanyang University, Seoul, Korea

Abstract

The Korean three-way stop contrast (lenis, aspirated, fortis) is currently undergoing a sound change, such that the primary cue distinguishing lenis and aspirated stops is shifting from voice onset time (VOT) to F0. Despite recent discussions of this shift, research on voice quality, traditionally considered an additional cue signaling the contrast, remains sparse. This study investigated the extent to which the associated voice quality [as reflected in the acoustic measurements of H1*–H2*, H1*–A1*, and cepstral peak prominence (CPP)] contributes to the three-way stop contrast, and how the realization is conditioned by prominence- vs. boundary-induced prosodic strengthening amid the ongoing sound change. Results for 12 native Korean speakers indicate that there was a substantial distinction in voice quality among the three stop categories with the breathiness of the vowel being the greatest after the lenis, intermediate after the aspirated, and least after the fortis stops, indicating the role of voice quality in the maintenance of the three-way stop contrast. Furthermore, prosodic strengthening has different effects on the contrast and contributes to the enhancement of the phonological contrast contingent on whether it is induced by prominence or boundary.

Keywords: Korean three-way stop contrast, voice quality, prosodic structure

1. Introduction

Korean has a three-way voiceless stop contrast between aspirated, fortis, and lenis stops (e.g., Cho et al., 2002; Kang, 2014). Traditionally, voice onset time (VOT), among others, has been regarded to be the primary feature, which distinguishes the three stop categories: fortis stop is produced with the shortest VOT, aspirated stop with the longest VOT, and lenis with an intermediate VOT. Other phonetic cues include the fundamental frequency (F0) and the voice quality of the vowel immediately following the stop.

Specifically, lenis stop has been considered to be associated with low F0, and aspirated and fortis stop to be associated with high F0. As for voice quality, vowel is known to be the most breathy after the lenis, intermediate after the aspirated, and the least breathy after the fortis stop (Cho et al., 2002).

However, recent studies have provided evidence that the lenis and aspirated stop contrast is now primarily cued by F0, with higher F0 for aspirated stops and lower F0 for lenis stops, especially in younger Koreans' speech (e.g., Bang et al., 2018; Kang, 2014). VOTs, which were previously the primary distinguishing cue between lenis and aspirated stops, have merged to the point where the contrast between the two categories is no longer present. Despite these recent findings and discussions, examination on voice quality and its relevance to the sound change remains notably absent in current literature.

Related to the shift in cue primacy (from VOT to F0), one account interprets the replacement of VOT with F0 as a sign of a 'tonogenetic' sound change in Seoul Korean (Bang et al., 2018; Kang, 2014). This account suggests that the phonetically driven low-level F0 differences have been exaggerated and integrated into the language's phonological distinctions, resulting in distinct tonal features (high F0 vs. low F0) in the segmental phonology of Seoul Korean. The significance of the segmental voicing feature (VOT) is attenuated, ultimately leading to a merger of VOTs. This account views the cue shift in Seoul Korean as mirroring the *transphonologization* observed in instances of 'tonogenesis' found in other languages such as Khmer and Afrikaans, where F0 transitions from a secondary phonetic property of a laryngeal contrast to a primary cue (Bang et al., 2018).

Aside from the transphonologization account, a new prosodic account has been put forward, arguing that the sound change in Seoul Korean is best understood as a prosodic-structurally conditioned variation in the utilization of the segmental voicing feature (i.e., VOT) versus the post-lexically available tones within the intonational phonology of the language (Choi et al., 2020). The intonational structure of Seoul Korean utilizes an accentual phrase (AP; above a prosodic word and below an Intonational Phrase in the prosodic hierarchy), which is assumed to be specified with a THLH tonal pattern (Jun, 1993). AP-initial tone (T) depends on the laryngeal feature of the segment: H tone assigned for aspirated or tense consonants, otherwise, L tone. Under this account, the cue shift takes place to the extent that the tones are available in certain prosodic contexts (i.e., phrase-initial positions) to reduce redundancy, and subsequently to the prominent prosodic context (i.e., under focus) where post-lexical tone is likely to be attracted.

The goal of the present study is to examine the voice quality differences associated with the three stops and explore whether these differences contribute to the three-way contrast amid the ongoing sound change. Additionally, motivated by the prosodic account proposed by Choi et al. (2020), we will also investigate how the voice quality difference may be conditioned by two prosodic-structural factors: focus-induced prominence and prosodic boundary, which have been found to interact with the recent sound change of the three-way stop contrast.

Several predictions can be formed, considering that prosodically prominent positions (i.e., under prominence and phrase-initially) are known to exhibit segmental realization in linguistically meaningful ways (Cho et al., 2017; Kim et al., 2018). One possibility is that the strengthening would be phonetically manifested by heightened glottalization across the board as sounds are known to be glottalized (or become creakier) in prosodic strengthening environments (Cho et al., 2017; Dilley et al., 1996; Garellek, 2014; Pierrehumbert & Talkin, 1992). Previous research on English has demonstrated that the degree of glottalization in vowel-initial words is greater under accent and/or in domain-initial positions (e.g., Dilley et al., 1996; Garellek, 2014; Pierrehumbert & Talkin, 1992). Similarly, Cho et al. (2018) illustrated that both prominence and prosodic boundary increase the degree of glottalization of initial vowels in South Kyungsang Korean. These findings align with a view that prosodic strengthening entails an increase in articulatory force which applies to both laryngeal and supralaryngeal articulation (Fougeron, 1999). Consequently, the prosodic strengthening effect may increase the laryngeal muscular tension, thereby increasing the degree of glottalization. If the three-way contrastive stops are produced with an increase in laryngeal muscular tension, the breathiness of the following vowel across the three-way stop categories would be reduced in prosodically prominent positions.

Alternatively, however, previous studies of prosodic strengthening have also indicated that prosodic strengthening extends beyond a mere low-level phonetic effect, making reference to the phonological system of a given language, enhancing the phonological contrast (Cho & Jun, 2000; Cho & McQueen, 2005; Kim et al., 2018). This leads to a predication that prosodic strengthening would increase the breathiness of the vowel after the lenis stop while the creakiness of the vowel is reinforced after the fortis stop, enhancing the three-way stop contrast.

To explore these possibilities, the present study investigates whether the three-way phonological contrast of word-initial stops manifest itself in the voice quality of the following vowel in Seoul Korean in prosodic strengthening environments. Spectral tilt has often been used to investigate voice quality across diverse languages (e.g., Garellek, 2022).

Thus, the amplitude differences between the first and second harmonics (H1*−H2*) and between the first harmonic and first formant (H1*−A1*) were taken as acoustic indexes of the voice quality with higher value indicating greater breathiness in the vowel (Garellek, 2013). In addition to spectral tilt, we include a noise measurement, as these two measurements (i.e., spectral tilt and noise measure) are often interpreted in tandem to distinguish the three-way distinction of breathy, model, and creaky voice (Garellek, 2019).

Thus, cepstral peak prominence (CPP), which detects noise components in an acoustic signal such as breathiness, was taken as another acoustic index of voice quality with lower value indicating greater breathiness in the vowel (Hillenbrand & Houde, 1996). It is worth noting that CPP values for creaky voice may be lower than that of modal voice, but they have been reported to distinguish creaky voice from breathy voice, which exhibits even lower CPP values (Esposito, 2012). Given that the degree of breathiness is often inversely correlated with the degree of creakiness (Cho et al., 2002; Gordon & Ladefoged, 2001), these measures will also be used to assess where in the breathy-creaky continuum the following vowel of each stop category falls.

2. Methods

2.1. Participants

Twelve native Seoul Korean speakers (7 females, 5 males)

participated in the study. They were all undergraduate students born and raised in Seoul/Gyeonggi Korea in their 20s (range, 21–28; mean, 24.7), who had limited (less than 2 years) overseas experience. All received financial compensation for their participation.

2.2. Speech Materials and Recording Procedure

Four triplets of Korean monosyllabic CVC syllables were used as target words (Table 1). Most of them were nonce words created to meet the following criteria for a larger corpus: The onset of the word was one of the three-way contrastive stops: lenis (/p, t/), aspirated (/pʰ, tʰ/), or fortis (/p*, t*/).

Table 1. List of target words by stop categories

Lenis	Aspirated	Fortis
박 /pak/	팍 /pʰak/	빡 /p*ak/
밧 /pat/	팟 /pʰat/	빳 /p*at/
답 /tap/	탑 /tʰap/	땁 /t*ap/
닷 /tat/	탓 /tʰat/	땃 /t*at/

The following vowel was controlled as /a/ and the coda consonant was either /k, t, p/. The target words were placed in carrier sentences with different prosodic renditions (Table 2). Prompt sentences (A in Table 2) were used to help forming a context of the mini-dialogue in which Speakers A and B were playing some kind of board game with cards. As for the boundary conditions to test domain-initial strengthening effects, the target word was placed either in phrase-initial position (IP-initial) or in phrase-medial position (IP-medial). For prominence, it either received focus by contrasting the onset consonant with /m/ (the focused condition) or unfocused by placing a focused element in another location in the sentence (the unfocused condition).

Speakers were asked to produce the test sentences in response to the prompt questions. Instead of the full written texts of stimuli sentences, visual clues for the carrier sentences were provided on a computer screen. For example, the screen showed two cards on which a monosyllabic test word was written on each of them in a contrastive way (e.g., *pak* vs. *mak*). The target word (e.g., *pak*) was marked with "O" and its contrasting word (e.g., *mak*) with "X". The pre-recorded voice, by two native speakers (1F, 1M) recorded prior to the experiment, was played through the loudspeaker, asking the speaker whether the next word to pick would be the contrasting word (marked with "X"). The speaker, cued by an "O" mark on the correct (target) word on the screen, was instructed to correct it by saying that the other one should be picked, thus making (corrective) focus on the target word. Given that the carrier sentences were simple, participants were able to produce the intended sentences in response after having received an about 10-minute training session. Acoustic data were collected in a soundproof booth using a Tascam HC-P2 digital recorder and a SHURE KSN44 condenser microphone at a sampling rate of 44 kHz at Hanyang Institute for Phonetics and Cognitive Sciences of Language.

In total, 2,304 tokens were collected (12 target words×2 boundary types×2 focus types×4 repetitions×12 speakers), and 2,037 tokens were used for further analysis, discarding tokens with unintended prosodic rendition checked by two trained Korean ToBI transcribers.

Table 2. Example sentences with the target word *pak* in different prosodic context. Focused words are in bold.

Conditions	Example sentences
IP-initial, Focused	A: [ipʌn tanʌnɯm **mak**satʃʲin twienonni]? 이번 단어는 **막**사진 뒤에 놓니? "This time, do I place the word (card) behind the picture of mak?" B: [ani]. IP [**pak**satʃʲin twi]. IP [twɛssʌ]? 아니. **박**사진 뒤. 됐어? "No. Behind the picture of pak. Got it?"
IP-initial, Unfocused	A: [ipʌn tanʌnɯm paksatʃʲin **apʰ**enonni]? 이번 단어는 박사진 **앞**에 놓니? "This time, do I place the word (card) in front of the picture of pak?" B: [ani]. IP [paksatʃʲin **twi**]. IP [twɛssʌ]? 아니. 박사진 **뒤**. 됐어? "No. **Behind** the picture of pak. Got it?"
IP-medial, Focused	A: [ipʌn tanʌnɯm ap*a **mak**satʃʲin twienonni]? 이번 단어는 아빠 **막**사진 뒤에 놓니? "This time, do I place the word (card) behind dad's picture of mak?" B: [ani]. IP [a*pa **pak**satʃʲin twi]. IP [twɛssʌ]? 아니. 아빠 **박**사진 뒤. 됐어? "No. Behind dad's picture of **pak**. Got it?"
IP-medial, Unfocused	A: [ipʌn tanʌnɯm ap*a pak.satʃʲin **apʰ**enonni]? 이번 단어는 아빠 박사진 **앞**에 놓니? "This time, do I place the word (card) in front of dad's picture of pak?" B: [ani]. IP [a*pa paksatʃʲin **twi**]. IP [twɛssʌ]? 아니. 아빠 박사진 **뒤**. 됐어? "No. **Behind** dad's picture of pak. Got it?"

2.3. Measurements

H1*–H2*, H1*–A1*, and CPP were measured as indexes of the degree of breathiness (or creakiness), obtained by VoiceSauce (Shue, 2010; Shue et al., 2011) (see Figure 1 for a schematization of spectral tilt measures). Note that * here indicates corrected measures for the effect of formant frequencies (Iseli & Alwan, 2004; Iseli et al., 2007). The values were obtained at the 25% and 50% points of the vowel.
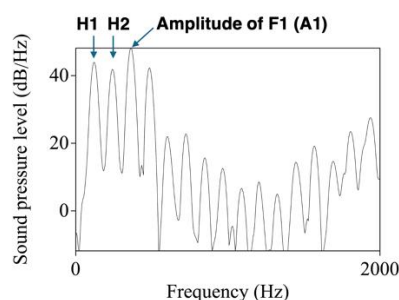

Figure 1. Schematization of spectral measurements of H1-H2 and H1-A1.

2.4. Statistical Analyses

The effects of Boundary and Focus on H1*–H2*, H1*–A1*, and CPP were examined by linear mixed-effects analysis using *lme4* package (Bates et al., 2015) in R (R Core Team, 2024). Time point factor (25%, 50%) was added to assess the extent to which the prosodic effects may vary (or be maintained) over the vowel. The factors of Focus, Boundary, and Timepoint were contrast coded, and the reference level for Stop factor was set as aspirated stop. Regarding the random structure, a maximal structure, as justified by

the design (Barr et al., 2013), incorporating by-speaker and by-item intercepts and slopes for the fixed effects, was employed, as long as the models converged. In cases where models failed to converge, random slopes with the least variances were eliminated. For models showing significant interactions between factors, pair-wise comparisons were conducted using the *emmeans* package with Tukey adjustments (Lenth, 2024). For the purpose of the present study, results that are directly related to the research questions (i.e., main effects of Stop and its interaction with Focus and Boundary) will be reported.

3. Results

Figure 2 summarizes the effect of Stop by timepoints in the vowel for H1*–H2*, H1*–A1*, and CPP. There was a significant main effect of Stop on H1*–H2* such that aspirated stop was *less* breathy than lenis stop (β=3.513, t=3.104, p<.01), but marginally *more* breathy than fortis stop (β=−1.361, t=−1.944, p=.074), suggesting that the difference on vowel quality among the three-way stop categories remains significant. There was a significant Stop× Timepoint interaction with fortis stop (β=−0.846, t=−2.663, p<.01), such that the difference between aspirated and fortis stop disappeared at later point in the vowel.

On H1*–A1*, there was a significant main effect of Stop with both fortis (β=−5.379, t=−5.202, p<.001) and lenis stop (β=6.033, t=5.834, p<.001), showing a three-way distinction among the stops. As shown in Figure 2 (b), vowels were *most* breathy (greatest H1*–A1*) after the lenis stop and *least* breathy (smallest H1*–A1*) after the fortis stop. H1*–A1* was intermediate for aspirated stops, showing a three-way stop contrast. There was no interaction between Stop×Timepoint, indicating that the difference between the three stops remained consistent across the two timepoints. H1*–A1* demonstrated the most clear distinction between the three categories.

Finally, there was a significant main effect of Stop on CPP such that aspirated stop showed the smallest CPP (most breathy) compared to both lenis stop (β=1.311, t=4.027, p<.001) and fortis stop (β=3.272, t=10.048, p<.001). Lenis stop was significantly *more* breathy than fortis stop, as shown in Figure 2 (c). The result suggests that the vowel quality difference also remains significant in CPP measure. There was a significant Stop×Timepoint interaction with fortis stop (β=−0.669, t=−3.196, p<.01), such that the difference between aspirated and fortis stop becomes greater at later point in the vowel.

In the following section, prosodic-structural effects on each voice quality measure (H1*–H2*, H1*–A1*, CPP) will be reported.
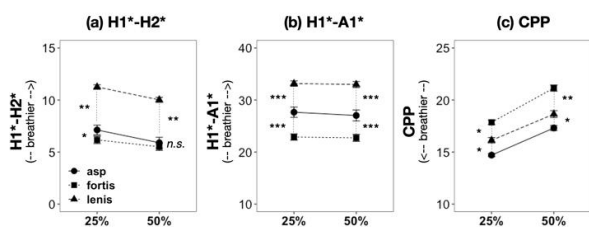
Figure 2. Effect of Stop on (a) H1*–H2*, (b) H1*–A1* and (c) CPP (*p<.05, **p<.01, ***p<.001). CPP, cepstral peak prominence.

3.1. H1*–H2*

Effect of prominence. There were significant interactions between Stop×Focus (fortis, β=−2.978, t=−9.241, p<.001; lenis, β=−1.703, t=−5.272, p<.001) as well as significant interactions between Stop×Focus×Timepoint (fortis, β=−2.308, t=−3.634, p<.001; lenis, β=−1.379, t=−2.151, p<.05) on H1*–H2*. The interaction was due to the fact that the focus effect was significant only at the 25% point after the aspirated stop. As illustrated in Figure 3 (a), H1*–H2* value for the aspirated stop was greater in the focused than in the unfocused condition at 25%, indicating that the breathiness of the aspirated stop was increased under focus. Another attribute to the interaction was the different effect sizes of Stop in the focused vs. unfocused conditions. That is, the difference between the stops was larger in the focused condition than in the unfocused condition (Figure 3 (b)).
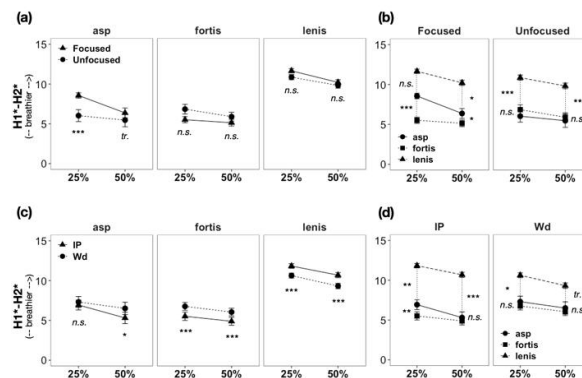
Figure 3. Effects of prosodic factors on H1*–H2*. The Stop×Focus interaction is illustrated in (a) by stop category and in (b) by focus condition. The Stop×Boundary interaction is illustrated in (c) by stop category and in (d) by boundary condition. Note that the difference between lenis and fortis stops was significant in all cases (*p<.05, **p<.01, ***p<.001).

Effect of boundary (domain-initial). There were significant interactions between Stop×Boundary (fortis, β=−2.308, t=−3.634, p<.001; lenis, β=−1.379, t=−2.151, p<.05). As shown in Figure 3 (c), the boundary effect was significant for the fortis and the lenis stops at both 25% and 50%, but only at 50% for the aspirated stop. Interestingly, the presence of a larger boundary had an opposite effect for [fortis, aspirated] vs. [lenis stop]: IP-initially, the vowels (compared to the IP-medial ones) showed smaller H1*–H2* (*less* breathy) after the fortis and aspirated stop, but larger H1*–H2* (*more* breathy) after the lenis stop. This finding aligns with the phonological contrast enhancement account (as discussed by Cho & McQueen, 2005; Kim et al., 2018), which predicted that prosodic strengthening would increase the breathiness of the vowel following the lenis stop, but increase the creakiness of the vowel following the fortis stop, thereby intensifying the three-way stop contrast. The interaction was also attributable to the fact that the effect of Stop was greater in the IP-initial position than in the IP-medial position.

Interaction between prominence and boundary. There were significant interactions between Stop×Focus×Boundary (fortis, β=−2.308, t=−3.634, p<.001; lenis, β=−1.379, t=−2.151, p<.05). As shown in Figure 4, the interaction stemmed in part due to the fact that the focus effect in the lenis stop (*more* breathy under focus) was only significant in the absence of an IP boundary. On the other hand, the boundary effect in the aspirated stop (*less* breathy IP-initially) was only significant in the unfocused condition. Another interesting point to note is that the difference between the three stop categories was the greatest in the extreme position of prosodic strengthening,

i.e., when both focused and IP-initial. Stop×Focus×Boundary interaction did not further interact with Timepoint.
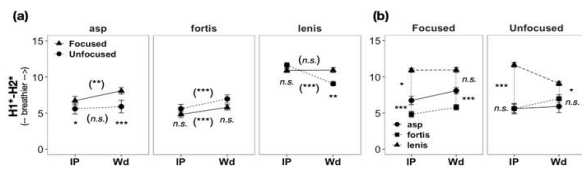

Figure 4. Focus×Boundary interaction on H1*–H2*. The effect of Boundary is illustrated in (a) by stop category and in (b) by focus condition. Asterisks in parentheses indicate pair-wise comparisons between IP vs. Wd conditions within each Focus condition. Note that the difference between lenis and fortis stops was significant in all cases (*p<.05, **p<.01, ***p<.001).

3.2. H1* – A1*

Effect of prominence. There were significant Stop×Focus interactions (fortis, β=−5.043, t=−8.908, p<.001; lenis, β=1.229, t=2.167, p<.05) on H1*–A1*. The interaction was due to the fact that the focus effect was significant in the vowel after the fortis and marginally significant after the lenis stop, but not after the aspirated stop (Figure 5 (a)). An opposite direction of focus effect was found for the fortis vs. the lenis stop, with the focus effect decreasing H1*–A1* for the fortis stop (i.e., *less* breathy under focus), but increasing H1*–A1* for the lenis stop (i.e., *more* breathy under focus), enhancing the three-way stop contrast under prominence. From a different perspective, the interaction was also in part due to the difference in the effect size of Stop in the focused vs. unfocused conditions. As shown in Figure 5 (b), the Stop effect was larger in the focused condition compared to unfocused condition.

Effect of boundary (*domain-initial*). There were Stop×Boundary interactions (fortis, β=1.181, t=2.087, p<.05; lenis, β=5.602, t=9.874, p<.001) on H1*–A1*, which was due to the boundary effect being significant after the lenis stop in both timepoints, but only at 50% after the aspirated stop. Vowels after the lenis stop were *more* breathy (greater H1*–A1*) in the IP-initial than in the IP-medial position, but less breathy (smaller H1*–A1*) at 50% after aspirated stop in the IP-initial than in the IP-medial position. Moreover, the effect of Stop was larger in the IP-initial position than in the IP-medial position (Figure 5 (d)), showing clear stop distinction in the domain-initial position.
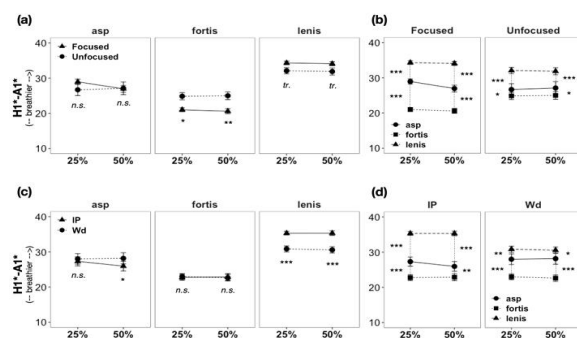

Figure 5. Effects of prosodic factors on H1*–A1*. The Stop×Focus×Timepoint interaction is illustrated in (a) by stop category and in (b) by focus condition. The Stop×Boundary×Timepoint interaction is illustrated in (c) by stop category and in (d) by boundary condition. Note that the difference between lenis and fortis stops was significant in all cases (*p<.05, **p<.01, ***p<.001).

Interaction between prominence and boundary. There was a significant interaction between Stop×Focus×Boundary with lenis stop (β=−9.012, t=−7.950, p<.001). As shown in Figure 6, the interaction stemmed in part due to the fact that, in the lenis stop, the focus effect (*more* breathy under focus) was only significant in the absence of an IP boundary, and the boundary effect (*more* breathy IP-initially) was only significant in the absence of focus. Conversely, the boundary effect in the aspirated stop (*less* breathy IP-initially) was only significant in the presence of focus. Stop×Focus×Boundary interaction did not further interact with Timepoint.
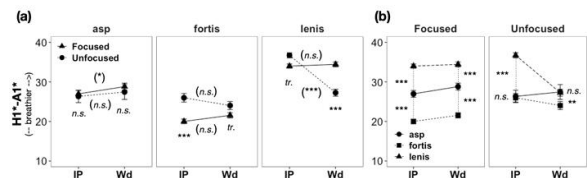

Figure 6. Focus×Boundary interaction on H1*–A1*. The effect of Boundary is illustrated in (a) by stop category and in (b) by focus condition. Asterisks in parentheses indicate pair-wise comparisons between IP vs. Wd conditions within each Focus condition. Note that the difference between lenis and fortis stops was significant in all cases (*p<.05, **p<.01, ***p<.001).

3.3. Cepstral Peak Prominence (CPP)

Effect of prominence. Significant interactions between Stop× Focus were detected (fortis, β=1.944, t=9.183, p<.001; lenis, β=−2.905, t=−13.699, p<.001). The interaction stemmed due to the focus effect being significant for fortis and lenis stops, but not for aspirated stop. However, the direction was opposite for the fortis vs. the lenis stop, with CPP increasing for the fortis stop (i.e., *less* breathy), but decreasing for the lenis stop (i.e., *more* breathy) under focus, as can be seen in Figure 7 (a). Again, this result supports the phonological contrast enhancement account (Cho & McQueen, 2005; Kim et al., 2018), which anticipates enhanced contrast under prosodic strengthening. There was also a significant Stop× Focus×Timepoint interaction with lenis stop (β=1.536, t=3.633, p<.001). The interaction was in part due to the focus effect in aspirated stop being significant at 25% but only marginally significant at 50% with the directions being opposite (Figure 7 (a)). From an alternate viewpoint, the interaction occurred because, while the differentiation among the three-way stops was delineated as [fortis, lenis] vs. [aspirated] in the unfocused condition, it shifted to [fortis] vs. [lenis, aspirated] in the focused condition, as shown in Figure 7 (b).

Effect of boundary (*domain-initial*). There was a significant interaction between Stop×Boundary with lenis stop (β=−2.562, t=−7.776, p<.001). As shown in Figure 7 (c), presence of an IP boundary significantly decreased CPP (*more* breathy) of the vowel after the lenis stop at both timepoints, whereas such effect was not found for the aspirated stop. Post-hoc analysis revealed that there was also similar effect of boundary for the fortis stop, but was significant only at 25% and marginally significant at 50% timepoint. Much like the result for the focus effect, from another angle, the distinction between the three-way stops was drawn as [fortis, lenis] vs. [aspirated] IP-medially, whereas the distinction was drawn as [fortis] vs. [lenis, aspirated] IP-initially, as illustrated in Figure 7 (d).
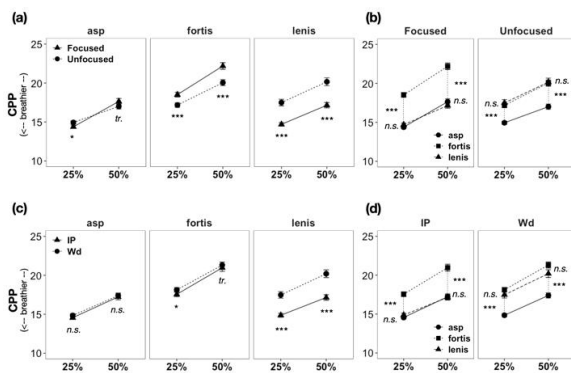
Figure 7. Effects of prosodic factors on CPP. The Stop×Focus interaction is illustrated in (a) by stop category and in (b) by focus condition. The Stop×Boundary interaction is illustrated in (c) by stop category and in (d) by boundary condition ($^*p<.05$, $^{**}p<.01$, $^{***}p<.001$). CPP, cepstral peak prominence.

Interaction between prominence and boundary. There was a significant Stop×Focus×Boundary interaction (lenis, β=4.598, t=10.845, $p<.001$). As shown in Figure 8, the interaction stemmed in part because CPP for the lenis stop was the greatest when IP-medial and unfocused, which is the context where lenis stop becomes voiced intervocalically. In another perspective, the interaction comes from the fact that the focus effects for the aspirated and lenis stop were only significant in the absence of an IP boundary, with focus *decreasing* CPP of the vowels. Similarly, the boundary effect was only significant in the unfocused condition, with IP boundary *decreasing* CPP, for all three stop categories. Stop×Focus×Boundary interaction did not further interact with Timepoint.
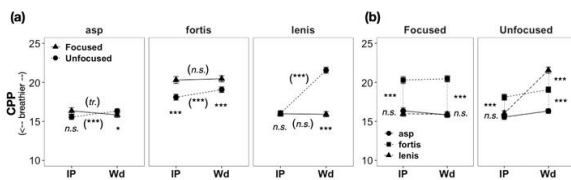

Figure 8. Focus×Boundary interaction on CPP. The effect of Boundary is illustrated in (a) by stop category and in (b) by focus condition. Asterisks in parentheses indicate pair-wise comparisons between IP vs. Wd conditions within each Focus condition. Note that the difference between lenis and fortis stops was significant in all cases ($^*p<.05$, $^{**}p<.01$, $^{***}p<.001$). CPP, cepstral peak prominence.

4. Discussion

One of the basic findings of the present study reveals that the difference in voice quality of the following vowel, as measured by acoustic parameters (H1*–H2*, H1*–A1*, and CPP), contributes to a three-way phonetic distinction among the Seoul Korean stop sounds produced by young speakers. H1*–A1* displayed the most clear three-way distinction among stop categories, while H1*–H2* primarily distinguished lenis stop from others, and CPP distinguished fortis stop from others. Notably, the amount of breathiness is largest for the lenis stop, intermediate for the aspirated stop, and smallest for the fortis stop, consistent with what was reported 22 years ago (Cho et al., 2002). This indicates that the voice quality difference has continued to underlie the three-way stop contrast. This finding is interesting, given the purported sound

change which has mainly been discussed regarding the shift between the two primary phonetic cues F0 and VOT (e.g., Bang et al., 2018; Kang, 2014) such that, for example, the difference between the lenis and the aspirated stops is signaled primarily by F0 with no difference in VOT. The results of the current study, however, demonstrate that the Korean stop contrast is still characterized by the laryngeal contrast at least at the phonetic level.

Another significant finding of the present study is that the three-way distinction in voice quality is further conditioned by prosodic strengthening factors: focus-induced prominence and boundary. The prosodic strengthening effects allow us to understand the phonological role of the phonetic distinction in voice quality. For instance, de Jong and colleague (de Jong, 2004; de Jong & Zawaydeh, 2002) propose that one way to assess the role of phonetic features in making phonological contrast may be to examine whether the phonetic feature participates in enhancing phonological contrast under focus-induced prominence. The results of the present study showed that the three-way stop contrast is indeed enhanced under focus, with the stops being substantially dispersed along the breathy-creaky phonetic continuum. The dispersion effect was observed in all measurements: H1*–H2*, H1*–A1*, and CPP.

Prosodic boundary (IP-initial vs. IP-medial) has also been found to influence the voice quality difference as a function of stop categories. As seen in the results section, the exact details of how the three-way voice quality distinction was modulated by boundary were somewhat different from the case of the prominence-driven strengthening effect. Boundary effect was mainly observed in H1*–H2*, whereas prominence effect was mainly observed in H1*–A1* and CPP, showing that prosodic strengthening may have different effects as a function of its source: prominence vs. boundary. This result further corroborates the findings of Peña et al. (2021), who illustrated that different types of glottalization, including both segmental and phrasal, show distinct acoustic properties. Likewise, prominence and boundary may also affect different acoustic characteristics of voice quality. Nonetheless, similar to the prominence effect, the boundary-related strengthening effect also induces an enhancement of the three-way stop contrast, demonstrating some degree of augmented dispersion of the stops along the breathy-creaky continuum (Cho & Jun, 2000).

Combining the results, the enhancement pattern under prosodic strengthening is that vowels become *more* creaky (less breathy) after the fortis stop, but *more* breathy (less creaky) after the lenis stop, contributing to the enhancement of the phonological contrast. Interestingly, the voice quality associated with the aspirated stop falls somewhere in between, which may be understood as an effort to retain the contrast by maintaining its intermediate position.

The results taken together imply that variation in the voice quality difference as a function of prosodic strengthening is not a mere low-level phonetic effect that would otherwise have applied to all three stops in a collective way, but is an outcome of the phonetic-prosody interface in reference to the phonological contrast in the language. This finding may be explained well with the prosodic account of the recent sound change in Korean three-way stop contrast, attributing the reduced VOT contrast to effort minimization (Choi et al., 2020). The prosodic account proposes that the existing post-lexical tones at phrase-initial positions (derived from the intonational structure) contribute to the segmental contrast, thereby diminishing the necessity of the redundant VOT

cue as a distinguishing feature. Our results on voice quality suggests that the three-way stop contrast is strengthened in the same positions, amplifying the redundancy of the VOT cue.

Finally, the results suggest that understanding the nature of laryngeal (voicing) contrast that occurs in Korean as well as in other languages requires multi-dimensional approaches to explore the phonetic realization of both the primary and other non-primary phonetic features (e.g., Al-Tamimi & Khattab, 2018; Kirby, 2018; Kong et al., 2011). It remains to be seen to what extent the voice quality difference is exploited by the listeners and how the voice quality cues may interact with F0 and VOT (cf., Kong et al., 2011).

5. Conclusion

This study explored the voice quality associated with the Korean three-way stop contrast (lenis, aspirated, fortis). While traditionally, the contrast between these stop categories has been primarily signaled by the interaction between VOT and F0, our findings shed light on the significant contribution of voice quality differences observable in the subsequent vowel (as reflected in H1*–H2*, H1*–A1*, and CPP). Moreover, we have examined how the realization of this contrast is conditioned by prominence- versus boundary-induced prosodic strengthening. The results from our analysis of twelve native Korean speakers highlight a notable distinction in voice quality among the three stop categories, emphasizing the crucial role of voice quality in maintaining the stop contrast. Additionally, our examination of prosodic strengthening reveals distinct effects on the contrast depending on whether the source originates from prominence or boundary. These findings offer valuable insights into the sound changes in Korean and have implications for further research in this area.

References

Al-Tamimi, J., & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, *71*, 306-325.

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.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Cho, T., & Jun, S. A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, *36*(1), 31-44.

Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193-228.

Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, *64*, 71-89.

Cho, T., Kim, D. J., & Kim, S. (2018). Prosodic strengthening in reference to the lexical pitch accent system in South Kyungsang Korean. *The Linguistic Review*, *36*(1), 85-115.

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*(2), 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. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*(4), 493-516.

de Jong, K., & Zawaydeh, B. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, *30*(1), 53-75.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*(4), 423-444.

Esposito, C. M. (2012). An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics*, *40*(3), 466-476.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: A review. *UCLA Working Papers in Phonetics*, *97*, 1-74.

Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation). University of California, Los Angeles, Los Angeles, CA.

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45,* 106-113.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *Routledge Handbook of Phonetics* (pp. 75-106). Oxford: Routledge.

Garellek, M. (2022). Theoretical achievements of phonetics in the 21st century: Phonetics of voice quality. *Journal of Phonetics*, *94*, 101155.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*(4), 383-406.

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, *39*(2), 311-321.

Iseli, M., & Alwan, A. (2004, August). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 669-672). Montreal, QC.

Iseli, M., Shue, Y. L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, *121*(4), 2283-2295.

Jang, J., Kim, S., & Cho, T. (2019). Prosodic-structural effects on voice quality associated with Korean three-way stop contrast. *Proceedings of the 19th International Congress of Phonetic*

Sciences (2416-2420). Melbourne, Australia.

Jun, S. A. (1993). *Prosodic structure and its interface with syntax in Korean* (Doctoral dissertation). Ohio State University, Columbus, OH.

Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics, 45*, 76-90.

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.

Kirby, J. P. (2018). Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages. *Journal of Phonetics, 71*, 326-354.

Kong, E. J., Beckman, M. E., & Edwards, J. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics, 39*(2), 196-211.

Lenth, R. V. (2024). emmeans: Estimated marginal means, aka least-squares means. R package (version 1.8.8). Retrieved from https://CRAN.R-project.org/package=emmeans

Peña, J., Davidson, L., & Orosco, S. (2021). The independence of phrasal creak and segmental glottalization in American English. *JASA Express Letters, 1*(7).

Pierrehumbert, J., & Talkin, D. (1992). 4 - Lenition of /h/ and glottal stop. In G. J. Docherty, & D. Robert Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90-117). Cambridge: Cambridge University Press.

R Core Team. (2024). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Shue, Y. L. (2010). The voice source in speech production: Data, analysis and models. Los Angeles: University of California. Retrieved from http://www.phonetics.ucla.edu/voiceproject/Publications/shue_dissertation.pdf

Shue, Y. L., Keating, P., Vicenik, C., & Kristine, Y. (2011, August). VoiceSauce: A program for voice analysis. *Proceedings of the 17th International Congress of Phonetic Sciences* (1946-1849). Hong Kong, China.

• **Jiyoung Jang**
  Post-doctoral Researcher, Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University
  222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
  Tel: +82-2-2220-2507
  Email: jiyoungljang@hanyang.ac.kr
  Fields of interest: phonetics, prosody

• **Sahyang Kim**
  Professor, Dept. of English Education, Hongik University
  94 Wausan-ro, Mapo-gu, Seoul 04066, Seoul, Korea
  Tel: +82-2-320-1848
  Email: sahyang@hongik.ac.kr
  Fields of interest: phonetics, prosody

• **Taehong Cho,** Corresponding author
  Director, Hanyang Institute for Phonetics and Cognitive Sciences of Language; Professor, Dept. of English Language and Literature, Hanyang University
  222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
  Tel: +82-2-2220-0746
  Email: tcho@hanyang.ac.kr
  Fields of interest: phonetics, laboratory phonolog