An acoustic and articulatory study on variation of high vowel devoicing across prosodic contexts and speakers in Korean

Jungyun Seo[1], Sahyang Kim[2,4] & Taehong Cho[3,4]
[1]*University of Michigan (USA),* [2]*Hongik University (Korea),* [3]*Hanyang University (Korea),* [4]*HIPCS*
sjungyun@umich.edu, sahyang@gmail.com, tcho@gmail.com

High vowels, particularly when surrounded by voiceless consonants, often undergo devoicing. This phenomenon can be attributed to physiological influences stemming from the tongue position during the production of high vowels [1]. Robust patterns of high vowel devoicing are observed in Japanese (e.g., [2]). In Japanese, high vowel devoicing is generally viewed as a phonological process whereby the [+voice] feature becomes disassociated and assimilated to the [-voice] feature of neighboring consonants (e.g., [3]), as evident in completely devoiced tokens that lack any trace of the vowel (e.g., [4]). In contrast, Korean devoicing is often described as a phonetic process, which can be considered to be due to the overlap between the glottis abduction gesture for voiceless consonants and the voicing gesture for the vowel ([5, 6]). Greater devoicing occurs when the overlap between these gestures is more extensive. Furthermore, the position of a vowel within a phrase has been found to exert influence on devoicing. For instance, completely devoiced tokens are more commonly found in phrase-initial positions as opposed to phrase-medial positions. This distinction can be accounted for by the larger overlap caused by the expanded glottis opening gesture at the domain-initial position, driven by the heightened [+spread glottis] feature due to domain-initial strengthening [7].

The present study aims to delve deeper into the nature of high vowel devoicing in Korean by examining the correlation between high vowel devoicing and tongue height. This investigation sheds new light on how this phonetic process relates to the physiological constraints imposed by tongue height. Furthermore, the study incorporates speaker variation to determine whether high vowel devoicing is an automatic process or one controlled by individual speakers. In doing so, the study explores variations in high vowel devoicing across different positions and in contexts where focus-induced prominence occurs. This exploration is important not only because devoicing patterns can be influenced by prosodic structural factors such as position and prominence but also because it enriches the contexts in which variation in high vowel devoicing can be observed.

An acoustic and articulatory study was conducted with 13 Seoul Korean speakers to investigate high vowel devoicing. The target words (p^hip^ha, p*ip*a) consisted of a high vowel /i/ surrounded by voiceless consonants, and they were produced in various prosodic structural contexts. Note that the same test word occurred in both IP-initial and IP-medial positions, as well as in focused and unfocused conditions. The focused condition involved contrasting the two target words, as presented in Table 1. Devoicing proportion was determined by calculating the ratio of acoustic duration between the voiceless portion and the entire syllable in the first syllable of CV.CV target words. Furthermore, tongue height (maxima during the vocalic movement) was measured using Electromagnetic Articulography (EMA) for the same set of target syllables. A total of 1483 tokens were examined, comprising 2 test words, 15 repetitions, 2 positions, 2 focus conditions, and 13 speakers.

The results revealed a gradient distribution of devoicing proportion (Fig. 1a), highlighting the phonetic nature of the devoicing process. Notably, completely devoiced tokens (20% of the data) were more prevalent in prosodically weak positions, indicating their susceptibility to coarticulatory effects from neighboring voiceless consonants. However, there was significant variability in devoicing patterns across speakers (Fig. 1b), suggesting individual differences in coarticulatory influences. Importantly, the correlation between devoicing proportion and tongue height was not clearly established ($\rho = 0.089$, $p < 0.001$), indicating that tongue height does not directly impact devoicing. This lack of correlation can be attributed to speaker variation. For instance, Fig. 2 illustrates speaker-specific effects of focus on tongue height and devoicing. Some speakers (Fig. 2c) exhibited increased tongue height with no effect on devoicing under focus, while others (Fig. 2b) showed unexpected patterns where heightened tongue position coincided with decreased (rather than increased) devoicing. These findings suggest that higher tongue position does not consistently induce more devoicing due to biomechanical factors. It appears that contrastive focus may enhance the [high] feature of the high vowel, potentially reinforcing the voicing feature. Yet, some speakers utilized both features, while others selectively suppressed devoicing under focus (Fig. 2a), and some did not demonstrate consistent modulation of tongue height

and devoicing in relation to featural enhancements. (Position effects also exhibited speaker variation not relying on tongue height, although it is not discussed in detail here due to space limitations.)

In conclusion, the findings suggest that high vowel devoicing in Korean exhibits a gradient phonetic process. However, the presence of speaker variation indicates that individual speakers adjust devoicing based on linguistic factors such as prosodic structure and phonological constraints, rather than relying solely on tongue height. Nevertheless, devoicing is more prevalent in prosodically weak positions, and certain speakers demonstrate a stronger inclination for devoicing. A broader implication is that if the number of speakers demonstrating robust devoicing increases, it has the potential to initiate sound changes, resulting in a more categorical devoicing process similar to what is observed in Japanese.

Table 1. Carrier sentences. Target words are underlined and a contrastive focus falls either on the target word or on the word that immediately follows it. Corrective contrast information was presented with the bold characters in the sentence.

Word	Boundary	Focus	Question sentence	Target sentence
피파 pʰipʰa	Phrase-initial	Focused	이번 단어는 삐빠뒤에 놔? [ipʌn tanʌnɯn p*ip*atwinoa?] Should I put the word behind p*ip*a this time?	아니야, **피파**뒤에 놔. [aniya.] [**pʰipʰa**twienoa.] No, put it behind **pʰipʰa**
		Unfocused	이번 단어는 피파 앞에 놔? [ipʌn tanʌnɯn pʰipʰaapenoa?] Should I put the word in front of pʰipʰa this time?	아니야, 피파**뒤**에 놔. [aniya.] [pʰipʰa**twie**noa.] No, put it **behind** pʰipʰa
	Phrase-medial	Focused	이번 단어는 누나 삐빠 뒤에 놔? [ipʌn tanʌnɯn nunap*ip*atwienoa?] Should I put the word behind sister's p*ip*a this time?	아니야, 누나**피파**뒤에. [aniya.] [nunap**pʰipʰa**twie.] No, put it behind sister's **pʰipʰa**.
		Unfocused	이번 단어는 누나 피파 앞에 놔? [ipʌn tanʌnɯn nunapʰipʰaapenoa?] Should I put the word in front of sister's pʰipʰa this time?	아니야, 누나피파**뒤**에. [aniya.] [nunapʰipʰa**twie**.] No, put it **behind** sister's pʰipʰa.

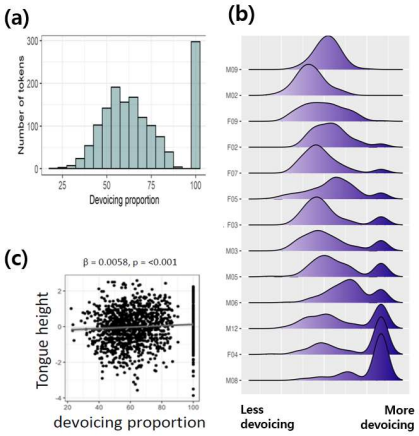

Fig.1 (a) Distribution of devoicing with 100 on the x axis indicating complete devoicing; (b) variation in distribution of devoicing across speakers ; (c) corelation between tongue height and devoicing.

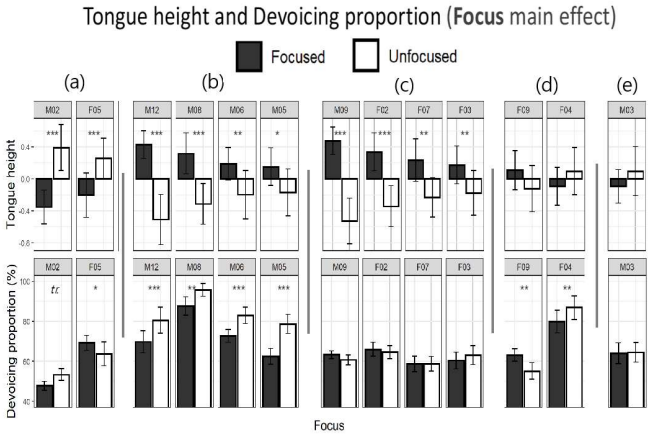

Fig.2 Prosodic modulation driven by Focus type in tongue height (upper panels) and devoicing proportion (lower panels) per speaker (***, **, * and tr. refer to $p < 0.001$, $p < 0.01$, $p < 0.05$, and $0.05 < p < 0.06$ in statistical analyses, respectively). Note that group (c) showed a higher tongue position in the focused condition compared to the unfocused condition without changing devoicing proportion; group (b) showed the focus effect only in the devoicing proportion ; and group (d) showed both a higher tongue position and less devoicing (more voicing) under focus.

References

[1] Ladefoged, P. (1973). The features of the larynx. *Journal of phonetics*, *1*(1), 73-83.

[2] Vance, T. J. (2008). The sounds of Japanese with audio CD. Cambridge University Press.

[3] Tsuchida, A. (2001). Japanese vowel devoicing: Cases of consecutive devoicing environments. *Journal of East Asian Linguistics*, 10(3), 225–245

[4] Beckman, M. E., & Shoji, A. (1984). Spectral and perceptual evidence for CV coarticulation in devoiced/si/and/syu/in Japanese. *Phonetica*, 41(2), 61–71.

[5] Jun, S. A., & Beckman, M. E. (1993). A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 67th annual meeting of the Linguistic Society of America, Los Angeles.

[6] Jun, S. A., & Beckman, M. E. (1994). Distribution of devoiced high vowels in Korean. In Third International Conference on Spoken Language Processing.

[7] Jun, S. A., Beckman, M. E., & Lee, H. J. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. UCLA Working Papers in Phonetics, 43–68.