DISTRIBUTION OF DEVOICED HIGH VOWELS IN KOREAN

Sun-Ah Jun

Linguistics, University of California, Los Angeles, California 90024-1543, USA

Mary E. Beckman

Linguistics, Ohio State University, Columbus, Ohio 43210-1298, USA

ABSTRACT

Korean has a process much like the well-studied "vowel devoicing" of Japanese. Evidence from Japanese suggests that the process is a phonetic undershoot of the vowel's gestures due to overlap with a preceding voiceless consonant. The similar Korean phenomenon would offer further support to this interpretation if vowel devoicing occurs differentially in the context of the three contrasting types of voiceless consonants, which have very different glottal gestures. The phonetic-undershoot account predicts that the vowel devoicing should be most common after plain fricatives or aspirated stops, where the glottal-opening gesture is largest and most likely to overweigh the vowel's voicing gesture. In syllables with lenis-stop initials, the vowel devoicing should occur less often in an accentual phrase medial position. The distribution of devoiced vowels was examined in a controlled corpus of dialogues containing CVCV target words, where the two consonants were voiceless and the first vowel was high. Each target word was produced either in phrase initial or in phrase medial position. The predictions were borne out. In addition, distributional differences involving the following consonant type suggest further testable differences about gestural overlap across syllable boundaries.

I. INTRODUCTION

In languages such as Japanese, syllables with short high vowels between voiceless consonants are often completely voiceless. This "devoicing" phenomenon has been noted for Montreal French [4] and Korean [8], but has been studied most extensively in standard (Tokyo) Japanese, where it has traditionally been described as a phonological rule that either categorically deletes the vowel or (the more typical characterization) categorically changes the vowel's phonological feature specification from [+voice] to [-voice]. A close examination of the phonetic evidence, however, shows that neither of these phonological characterizations is supported unambiguously. Indeed, all of the evidence is at least equally compatible with an alternative account. Perhaps the phenomenon is not the categorical result of a phonological rule at all, but instead is a phonetic undershoot of the vowel's gestures, involving overlap and blending with the consonant's gestures. In other words, suppose that vowel devoicing is another example of the kind of gradient phonetic effect resulting from more or less subtle adjustments to the magnitude and timing of otherwise invariant gestural specifications, as in Browman and Goldstein's account of casual speech processes in English [3].

Such a gestural overlap account is supported by the spectral patterns typically observed in devoiced vowels. In acoustic studies of the phenomenon, waveforms and spectrograms of target syllables show a continuum of degrees of devoicing, from tokens with one or two reduced glottal pulses to tokens having no trace even of the vowel's lingual gesture. This is true of studies of Montreal French [4] and Korean [8], as well as of Japanese [10][13]. A study of the perception of such voiceless syllables in Japanese [1] also supports the alternative account: minimal pairs such as [kaʃi] 'sweets' versus [kaʃu] 'poet' are distinguished much less accurately when the target syllable is voiceless, but still considerably better than at chance level, at about the same rate as *she* and *shoe* are distinguished in English when the fricatives are excised from context in experiments on coarticulation [15].

The gestural overlap account also implies that devoicing should be more commonly observed in just those conditions where the neighboring consonant's gestures overlap more with the vowel's. Thus further evidence for the account comes from examining the distributional characteristics of the phenomenon. Earlier studies have shown that the phenomenon occurs to varying extent for different segmental sequences and in different prosodic positions. For example, in Montreal French, the vowel is much less likely to devoice in the second-to-last syllable in a rhythm group, and never devoices in the absolute rhythm-group final syllable [3, 4, 12]. Both of these are positions where the vowel is prosodically lengthened and less likely to be completely covered over by the preceding consonant's gesture [12]. In Japanese, similarly, devoicing is limited to phonemically short vowels, and it most commonly affects the two high vowels, which are phonetically very short even when they are not devoised. Devoicing does occur sometimes with mid and low vowels [10][13], but here it is relatively rarer.

Since Korean contrasts three types of voiceless consonants that are known to have very different glottal gestures, distributional patterns for the similar Korean phenomenon should be even more revealing. For example, the top right-hand panel in Figure 1 shows Kagaya's [9] data on average glottal aperture for the three stop types in initial position. The average time traces are aligned at the stop's oral release (the vertical line in the figure) to emphasize the gestural differences. The aspirated stop has the largest glottal abduction, which is timed so that peak opening occurs around the oral release.

The work reported here was supported in part by the National Science Foundation through grant No. IRI-8858109.

The lenis stop shows a medium-sized abduction gesture ending shortly after the oral release. Finally, the fortis stop shows a much smaller abduction ending in a tight glottal seal just before the oral release. The gestural overlap account predicts that, in initial syllables, vowel devoicing should be most common after aspirated stops, where the consonant's glottal-opening gesture is largest and extends most into the vowel, and thus is most likely to induce complete undershoot of the vowel's voicing gesture. Conversely, devoicing should be least common after fortis stops, where the glottis is closed already before the consonant's oral release.

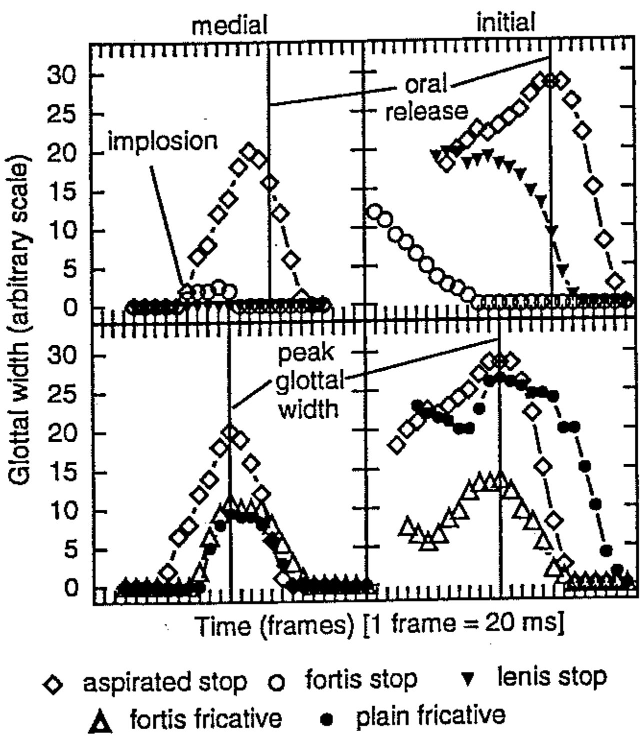

Fig. 1. Glottal aperture data, adapted from Kagaya [9]. Averaged time series for (top panels) aspirated versus fortis versus lenis stops, and for (bottom panels) fortis versus plain fricatives (versus aspirated stops shown again for comparison). Alignment of traces is at oral release of the stop (top) or at peak glottal opening (bottom), with medial phrase position on the left, and initial on the right.

This paper reports on the distribution of devoiced high vowels in a controlled corpus of dialogues which put target syllables in the context of different preceding and following consonants in different prosodic positions. The resulting patterns are interpreted in terms of what is known about the contrasting consonants' articulations. The above predictions concerning different preceding stop contexts were borne out unambiguously. In other contexts, the predictions were less clear, because the articulatory data in the literature is less complete or less clearly indicates the underlying gestural organization. For example, whereas a preceding fortis stop induced less devoicing than the corresponding fricative, a following fortis stop induced more. The differing degrees of glottal opening in the two consonant types (left hand panels in Fig. 1) would predict the first of these results but does not seem compatible with the second. On the other hand, this result may reflect other differences which interact with the airflow requirements for sustained voicing. For example, the oral gesture for a fortis stop may involve a faster, more ballistic closing movement than that into the corresponding sibilant. The faster movement, then, could result in a more immediate buildup of oral air pressure to impede voicing. Thus, the vowel devoicing pattern due to the following consonant type suggests further testable differences in gestural overlap across syllable boundaries.

II. METHODS

The corpus consisted of many disyllabic CVCV target words embedded in the dialogue shown in Figure 2. Since the dialogue contained two pairs of question and answer sentences, it provided tokens of two different target words in each of two different focus positions. In the Korean prosodic system, the most salient indication of focus in an utterance is the phrasing. A focused word is always initial to an **accentual phrase**, a tonally defined prosodic constituent that is associated with many phonetic processes such as voicing of phrase-medial lenis stops [6, 7]. All words after the focused word until the next major intonational boundary are then grouped together into the same accentual phrase. Thus, in our corpus of dialogues, the target word is produced first under focus, making its first consonant initial to an accentual phrase, and then is repeated with focus on the preceding deictic pronoun, so that the initial consonant of the target word is medial to its phrase.

(1) Q: ike kutuni?	"Is this a SHOE?"
(2) A: ani, tʃəke kutuja.	"No, THAT is a shoe."
(3) Q: kɨləm, ike tʰusuni?	"Then, is this a PITCHER?"
(4) A: iŋ, kike tʰusuja.	"Yes, THAT is a pitcher."

Figure 2. Sample dialogue. The underlined parts are focused and /kutu/ 'a shoe' and /tʰusu/ 'a pitcher' are the two target words. In sentences (1) and (3), the target is focused, and hence initial to the accentual phrase, whereas in (2) and (4), the target is post-focal and hence medial to the focused accentual phrase starting at /tʃəke/ or /kike/ 'that'.

There were 25 different target words, providing a balanced set of preceding and following consonant types for the first vowel, which was the segment of interest. This vowel was any one of the three high vowels of Korean, /i, u, ɨ/, and each of two surrounding consonants was one of the five voiceless consonant types -- aspirated, fortis, or lenis stop, or fortis or plain fricative. We devised 125 dialogues containing 500 tokens where 250 tokens are phrase initial and 250 tokens are phrase medial. For each phrasal condition, 25 target words were repeated ten times and randomized. (One target word -- [s'ukʰi] -- is a nonsense word, but speakers had no difficulty pronouncing it.) The dialogues were recorded in a sound-attenuated booth by each of six speakers, who read both the Q and A parts. There were three speakers of Seoul Korean and three of Chonnam Korean, one female and two males for each dialect.

We examined the status of voicing by displaying the portion of the utterance around each target word as a wide band spectrogram on a Kay DSP 5500 SonaGraph. Although the speakers all read the dialogues at a fairly constant rate in a well-modulated reading style, there was much variation in the voicing status of tokens even within the productions by each individual speaker, ranging from an obviously voiced type with a duration of up to 80 ms and clearly defined higher formants, to an extremely reduced type with no evidence of any glottal pulses between the spectral hallmarks of the first and second consonant. That is, the degree of voicing was quite gradient. In order to calculate the frequency of devoicing, we quantized this continuum into three categories: a vowel was "voiced" when the periodic energy from the laryngeal source was large enough for a long enough interval to visibly excite the second or higher formants on the spectrogram, "partially devoiced" when there were one or two weak glottal pulses visible at the bottom of the spectrogram, and "completely devoiced" when there was no visible glottal pulse for a vowel.

III. RESULTS AND DISCUSSION

Figure 3 shows the cumulative percentage of completely and partially devoiced vowels in the different contexts, averaged over the productions by all six speakers. As we predicted, when a phrase-initial consonant precedes the vowel (top panel), consonants with larger and longer glottal openings are generally associated with more instances of devoicing on the following vowel. A loglinear analysis of these frequencies showed significant main effects of consonant manner (stop versus fricative) and of phonation type (aspirated versus lenis versus fortis), and also a significant interaction. Post-hoc χ^2 tests on pairs with the most similar distributions showed the following significant differences: [s] > [tʰ] > [s'] = [t] > [t'] (Here we use the dental as the representative of the group -- e.g. [t'] stands for all fortis stops.) Thus, the fortis stop induces significantly less devoicing on the following vowel than a lenis stop, in keeping with the glottal aperture data in the literature, as illustrated in the top right-hand panel of Fig. 1 above. Also, the largest incidence of devoicing occurred after the non-fortis fricative and the aspirated stop, with somewhat more after fricative, in keeping with the fact that the glottis opens very wide in both of these consonant types, but stays at its peak opening longer in the fricative (see the bottom right-hand panel of Fig. 1).

A second major result evident in Fig. 3 is that there was a very large effect of the preceding consonant's phrasal position in the case of the lenis stop context (χ^2=58.4, $p<0.001$). None of the other consonantal contexts is associated with such a large difference between the two phrasal positions, and other than the reordering of the lenis stop context, the effect of preceding consonant in phrase-medial position was identical to that for the preceding consonant in phrase-initial position: [s] > [tʰ] > [s'] > [t] = [t']. This result also is unsurprising, given what we know about the accentual phrase. Although there are no studies contrasting glottal width in different positions in the accentual phrase, work by Jun [6, 7] and others [14] shows that when a word-initial lenis stop is medial to its accentual phrase, it typically shows voicing throughout much of its closure. This suggests that the glottal aperture patterns shown in Fig. 1 above for Kagaya's word-medial versus utterance-initial lenis stops actually reflect the effects of being medial versus initial to the accentual phrase. (Kagaya's corpus contrasted CV versus VCV nonsense words.) The aspirated and fortis stops and the two types of fricatives in Fig. 1 do not show any such qualitative difference in glottal aperture pattern.

The last panel in Fig. 3 shows the effect of the following (word-medial) consonant. There are three main results. First, the range of differences here is much less than in the top two panels. This suggests a less close association between the gestures of the vowel and the following consonant, in keeping with the fact that here the consonant belongs to the following syllable.

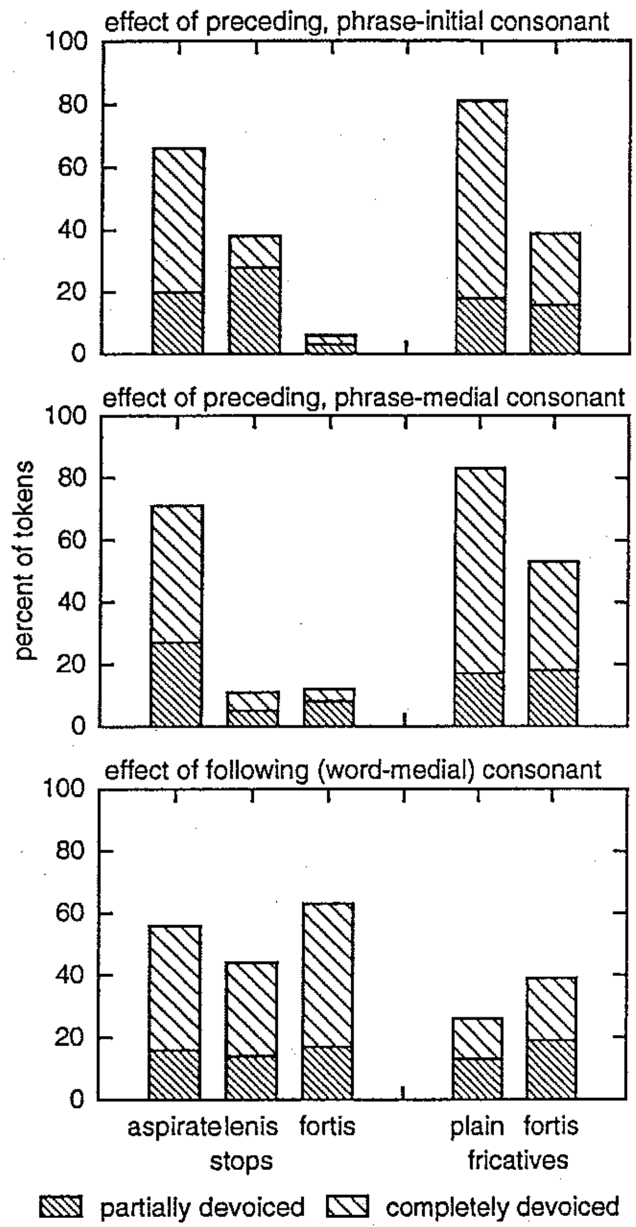

Figure 3. Percentage of tokens in which the vowel is fully and partially devoiced in the context of (top) different preceding consonants in phrase-initial position, (middle) different preceding consonants in phrase-medial position, and (bottom) different following consonants. The total is about 300 tokens (50 tokens x 6 speakers) for each of the bars in the first two panels, and 600 tokens for each of the bars in the last panel.

Second, a following stop was associated with more devoicing than a following fricative. This was the opposite of the pattern shown for the preceding consonant context, where the plain fricative induced the most devoicing, and the fortis fricative induced considerably more devoicing than the fortis stop. The greater incidence of devoicing induced by a following aspirated stop is not so surprising, since peak glottal opening is more reduced in the plain word-medial fricative than in the aspirated stop. However, the effect of a following fortis or lenis stop is not as easy to interpret in terms of the glottal aperture patterns in Fig. 1. That is, glottal aperture data in Fig. 1 cannot explain why lenis and fortis stops also induced more devoicing than the corresponding plain and fortis fricatives, since word-medially, the fricatives have larger glottal opening. On the other hand, kinematic studies of oral articulations often show faster, more ballistic movements into stops than into sibilant fricatives (e.g., [11]). Given the transglottal airflow requirements for voicing, and the quicker buildup of oral air pressure from the closure into the stop, glottal vibration may cease sooner in a vowel-stop sequence than a vowel-fricative sequence. Or, if the vowel is short and has a narrow oral constriction, the vowel's voicing gesture before a stop may not produce a glottal gesture of enough amplitude to show up on the spectrographic display.

The third result evident in the bottom panel of Fig. 3 is that a following fortis consonant was always associated with more vowel devoicing than the corresponding lenis stop or plain fricative. This result also differed qualitatively from the pattern for the preceding consonant, where the fortis consonants typically induced devoicing of the vowel in the fewest tokens. Given the virtually identical glottal aperture patterns in the lower left-hand panel of Fig. 1, the pattern for the fricatives is particularly puzzling. Data on EMG activity of the laryngeal muscles may be more revealing here. Average time traces of vocalis activity in VCV sequences for the five different Korean stops [5] show considerable differences between the effect of the consonants on preceding and following vowels. In the vowel following the consonant, there is less activity after a plain fricative than after a fortis fricative, suggesting a less pressed or breathier mode of phonation. In the vowel preceding the consonant, by contrast, there is less activity before a fortis fricative. Moreover, the difference between the vowels before the contrasting fricative types begins to emerge already at the release of the preceding consonant in the frame, suggesting an extreme overlap or an active anticipation of the devoicing gesture for the fortis fricative.

IV. CONCLUSION

We examined the distribution of devoiced high vowels in Korean and found that high vowels were more likely to be devoiced after consonants with a large glottal opening gesture and less likely to be devoiced after consonants with a small glottal opening gesture. We also found that the degree of glottal opening alone cannot explain the frequency of vowel devoicing with respect to the following consonant. Vowels were more likely to be devoiced before stops than before fricatives and more likely to be devoiced before fortis than before lenis or plain consonants. To explain these patterns, it seems necessary to consider not just the timing of the glottal opening gesture in the following consonant relative to the vowel's glottal closing gesture, but also such things as vocal fold tension and the relative oral and subglottal air pressures at the end of the vowel.

At the same time, none of these considerations are available to an account of the phenomenon that characterizes the devoicing as a categorical phnological rule. Thus, the evidence from Korean supports our proposal based originally on the phonetic data from Japanese. These vowel devoicing processes are much easier to understand if we represent them as the result of extreme overlap and submerging of the vowel's glottal gesture by the neighboring consonant's gestures.

REFERENCES

[1] M. E. Beckman, A. Shoji, "Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese," *Phonetica* Vol. 41, pp. 61-71, 1984.

[2] C. P. Browman, L. Goldstein, "Tiers in articulatory phonology, with some implications for casual speech," In Kingston & Beckman, eds., *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, pp. 341-376, 1990.

[3] H. Cedergren, Linguistics Department Speaker Series, Ohio State University, 1994.

[4] H. J. Cedergren, L. Simoneau, "La chute des voyelles hautes en français de Montreal: 'As-tu entendu la belle syncope?'" In Lemieux & Cedergren, eds., *Les tendences dynamiques du français parle a Montreal*, pp. 57-144, 1985.

[5] H. Hirose, H. S. Park, & M. Sawashima, "Activity of the thyroarytenoid muscle in the production of Korean stops and fricatives," *Ann. Bull. RILP*, No. 17, pp. 73-81, 1983.

[6] S.-A. Jun, "The accentual pattern and prosody of Chonnam Dialect of Korean," In Kuno et al., eds., *Harvard Studies in Korean Linguistics* III, 1989.

[7] S.-A. Jun, *The Phonetics and Phonology of Korean Prosody*, Ph.D. dissertation, the Ohio State University, 1993.

[8] S.-A. Jun, M. E. Beckman, "A gestural-overlap analysis of vowel devoicing in Japanese and Korean," Paper presented at the 67th Annual Meeting of the Linguistic Society of America, 1993.

[9] R. Kagaya, "A Fiberscopic and Acoustic study of the Korean stops, affricates and fricatives," *J. Phonetics*, Vol. 2, pp. 161-180, 1974.

[10] M. Kondo, "The effect of blocking factors and constraints on consecutive vowel devoicing in Standard Japanese," Fourth Conference on Laboratory Phonology, 1993.

[11] D. P. Kuehn & K. L. Moll, "A cineradiographic study of VC and CV articulatory velocities," *J. Phonetics*, Vol. 4, pp. 303-320, 1976.

[12] L. Levac, H. J., Cedergren, H. Perreault, "Phonetic evidence of narrow and wide temporal scope for prosodic constituents in French," *Proceedings of an ESCA Workshop on Prosody*, Working Papers, Department of Linguistics and Phonetics, Lund, No. 41, pp. 54-57, 1993.

[13] K. Maekawa, "Hatsuwa sokudo ni yoru yuusei kukan no hendoo," *Technical Report* SP89-148, Densi Zyoohoo Tuusin Gakkai, 1990.

[14] D. J. Silva, D. *The Phonetics and Phonology of Stop Lenition in Korean*. Ph.D. dissertation, Cornell University, 1992.

[15] G. H. Yeni-Komshian & S. D. Soli, "Recognition of vowels from information in fricatives: perceptual evidence of fricative-vowel coarticulation," *J. Acoust. Soc. Am.* Vol. 70, pp. 966-975, 1981.