TEMPORAL, SPECTRAL EVIDENCE OF DEVOICED VOWELS IN KOREAN

Yoonsook Mo

University of Illinois at Urbana-Champaign
ymo@uiuc.edu

ABSTRACT

Vowel devoicing is a phenomenon that is reported to occur in many languages such as Japanese, Parisian and Montreal French, Turkish and English. This paper investigates vowel devoicing in Korean. A devoiced vowel does not exhibit characteristic vocal tract resonances, and instead is realized as a long interval of aspiration or frication following consonant release, resulting in non-distinct segment boundaries between devoiced vowels and adjacent voiceless consonants.

This paper examines temporal and spectral evidence of devoiced vowels and, among other findings, reveals that in Korean devoiced high vowels are not segmentally deleted but phonetically masked, suggesting that vowel devoicing results from the overlap of glottal gestures. This paper also examines the effect of the preceding consonant place and manner, and the height and front/backness of vowels on devoicing.

Keywords: vowel devoicing, F2, gestural overlap

1. INTRODUCTION

Vowel devoicing has been reported for various languages [1, 3]. High vowels and schwa, which are intrinsically short in duration and more restricted in size of oral cavity, are often found to be devoiced or deleted in certain consonantal environments. There are two distinct approaches to the analysis of vowel devoicing or deletion. In earlier treatments of Japanese, high vowels in devoicing contexts are described as phonologically deleted [1]. In his electromyographic study, Hirose [2] observed that the glottis remains open throughout the production of voiceless vowels, suggesting that vowel devoicing is controlled by a higher neural process, rather than by a lower physiological process.

A more recent approach to vowel devoicing treats vowel devoicing as the result of physiological processes. Jannedy [3] and Jun and her colleagues [4, 5] argued that vowel devoicing in Turkish and in Korean results from gestural overlap of vowels with their adjacent voiceless consonants. That is, their argument is that glottis closing gestures for high vowels are overlapped with or delayed by surrounding glottis opening gestures for voiceless consonants, resulting in the devoicing of high vowels. Yoshioka [6] also observes that even if the glottis remains open throughout the production of voiceless vowels, muscular activity is unusually variable during the production of voiceless vowels. Davidson [7] claimed that English schwa elision can also be best described with the account of gestural overlap.

Previous works provide little in the way of a detailed acoustic description of devoiced vowels. Kondo [8] measures the average moraic durations of devoiced and voiced tokens in Japanese. Davidson [7] and Manuel et al. [9] discussed the acoustic outcomes of schwa reduction in English. The present paper reports the temporal and spectral measurements as well as the frequency measurements of devoiced vowels to answer the following research questions: (1) whether devoiced vowels in Korean are segmentally deleted or phonetically masked; (2) if so, whether the identity of high vowels is acoustically maintained in devoicing contexts; (3) what characteristics of the preceding consonants affect vowel devoicing. Results appear to support a gestural overlap interpretation of vowel devoicing.

2. MATERIALS AND MEASUREMENTS

2.1. Materials and Recording

Four female speakers of standard Seoul Korean aged 29 to 38 participated in this experiment. The materials are composed of 98 simple sentences which contain a disyllabic word, CVta embedded in a carrier sentence. The initial voiceless C of the disyllabic word is drawn from /k, k’, k^h, t, t’, t^h, p, p’, p^h, s, s’, tʃ, tʃ’, $tʃ^h$/ and the following target V is one of 7 monophthongs /a, æ, ə, ɨ, i, o, u/, with all combinations of initial C and V. The recordings were made using Praat or CSL.

2.2. Measurements

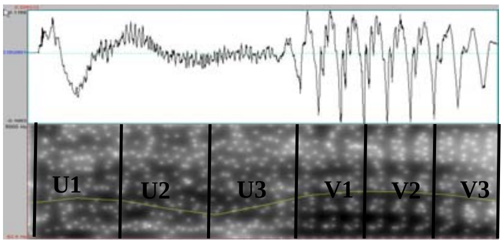
2.2.1 Durational measurement

Durations were measured from the onset of oral release or aperiodic frication to the offset of frication or aspiration of the target syllable. For tokens with appreciable voicing of the target vowel, durations from the onset of voicing to the offset of the vowel were also measured. Then tokens are divided into two groups. If there is no appreciable voicing of the vowel, they are called as "devoiced"; otherwise, they are called as "voiced".

2.2.1 Formant measurement

Averaged formant values, F1, F2, and F3, were measured in three separate regions based on the LPC analysis within a 25ms window, with a .0625ms time step in Praat. For tokens with no evidence of voicing energy, the devoiced syllable is divided into three equal intervals from the onset to the offset of the syllable. When there is some region of voicing in the target vowel, the consonantal and vowel regions are divided into three equal intervals respectively as shown in the following example.

Figure 1: The averaged formants are measured at three different intervals for devoiced (U1, U2, U3) and voiced regions (V1, V2, V3) respectively as illustrated with a syllable /pa/ in the spectrogram.



3. RESULTS

3.1. Effects of vowel height and backness

Overall, 104 tokens out of 1028 vowels in the target syllable (10.12%) are devoiced. However, separating high vowels from non-high vowels, 102 tokens out of 455 high vowels (22.4%) are devoiced while only 2 tokens out of 573 non-high vowels (0.35%) are devoiced. A non-parametric Pearson's χ^2 test shows statistical significance of effects of vowel height, (χ^2=122.09 > 6.635, α = 0.001). In other words, high vowels are much more likely to be devoiced than non-high vowels. Comparing the front/backness of target vowels, a Pearson's χ^2 test does not reveal a significant difference, (χ^2=0.916 < 6.635, α = 0.001). The front/backness of a target vowel does not show any influence on devoicing of that vowel.

3.2. Effects of the preceding consonants

3.2.1 Effects of the manner of articulation

In order to balance the data in terms of the degree of aspiration and the place of articulation of the preceding consonants, only the data preceded by a lenis, coronal consonant are analyzed. A Pearson's χ^2 test reveals that the manner of articulation of the preceding consonants is an important factor contributing to devoicing of the following vowels (χ^2=21.454 > 6.635, α = 0.001). A lenis fricative is more likely to devoice the following vowels (74.3%, 26/35) than a lenis affricate (27.3%, 9/33) or a lenis stop (8.6%, 3/35).

3.2.2 Effects of the degree of aspiration

Korean has a three way distinction in terms of the degree of aspiration: aspirated, lenis (slightly aspirated), and fortis (unaspirated, laryngealized). Three Pearson's χ^2 tests are performed to examine the effects of the degree of aspiration of the preceding consonants on devoicing of the following vowels. All of them show a significant effect of the degree of aspiration (χ^2=38.160 > 6.635 for affricates, 12.506 > 6.635 for stops, 24.633 > 6.635 for affricates and stops, α = 0.001). An aspirated affricate (88.6%, 31/35) tends to devoice a following vowel more frequently than a lenis (27.3%, 9/33) or fortis affricate (0%, 0/34). Like affricates, aspirated stops (27.6%, 27/98) are more likely to make a following vowel devoiced than lenis (4.9%, 5/102) or fortis (0.0%, 0/87) stops. Taking into account affricates and stops together, we find that the highly aspirated consonants are more likely to devoice the following vowels.

3.3. Durational effects

The mean difference between the duration of the voiceless portion and the entire syllable is compared for syllables with both voiced and devoiced high vowels following aspirated stops.Target syllables with an aspirated stop onset provide a balanced number of voiced and devoiced tokens of the target high vowels, whereas syllables with an aspirated affricate or a lenis fricative onset most often exhibit high vowel devoicing, while syllables with a fortis consonant onset exhibit no devoicing at all.

Due to the variation in intrinsic VOT durations of the aspirated stops depending on their place of articulation, three separate one-way ANOVA tests are performed for each preceding consonant, /p, t, k/. The durations of the front, central, and back high vowels are grouped together because they are found not to be significantly different from each other in likelihood of devoicing (p = .154).

Table 1: The mean syllable durations of devoiced vs. voiced tokens by the preceding consonants

Mean durations (ms)	k	p	t
voiced	87.97	83.11	84.40
devoiced	87.26	68.56	74.92

A one-way ANOVA shows that the mean syllable durations of devoiced tokens are not significantly different from those of voiced tokens (p = .203 for /p/, .062 for /t/, .809 for /k/). On the contrary, the mean durations of the voiceless interval of each target syllable from devoiced tokens are significantly longer than those of voiced tokens (p = .048 for /p/, .001 for /t/, .006 for /k/). In summary, even if vowels are devoiced, the syllable durations of both devoiced and voiced tokens remain the same.

Table 2: The mean durations of VOT of devoiced vs. voiced tokens by the preceding consonants

Mean durations (ms)	k	p	t
voiced	68.28	61.57	56.12
devoiced	87.23	68.56	74.92

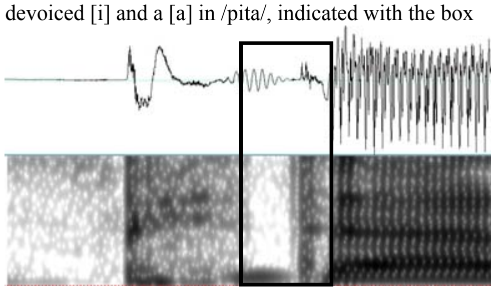
3.4. Spectral evidence

All high vowels are equally likely to be devoiced regardless of the front/backness. We measured the second formants of the target vowels in order to inspect the vowel trace within the devoiced region. Then, we compared F2 values of devoiced tokens according to the front/backness of the target vowels.

The results reveal that the mean F2s of devoiced tokens are consistently higher than those of voiced ones. Also, the mean F2s of devoiced tokens are significantly different from each other in relation to the front/backness of target vowels (p < .001 for /s/, p < .001 for /p/, .013 for /t/, p < .001 for /k/) except those with the aspirated affricate, which are marginally significant (p = .063).

In addition to F2s, the intervocalic voicing of /t/ is spectrally observed between a devoiced vowel and the canonical vowel, /a/. All Korean consonants are voiceless but often voiced intervocalically. Among four subjects, one subject consistently produced voicing of /t/ between a target vowel and /a/ in CVta. However, in some cases where a target vowel is completely devoiced and, therefore, the condition of intervocalic voicing of /t/ has not met, the regular glottal pulses (voice bars) of the voiced /t/ are observed shown in Figure 2. This indicates that the devoiced vowel retains its phonological voicing gesture in order to condition intervocalic voicing of the following consonant, /t/.

Figure 2: The intervocalically voiced /t/ between a devoiced [i] and a [a] in /pita/, indicated with the box



4. DISCUSSION

Like other languages, the height of a target vowel plays the most important role in vowel devoicing in Korean. Non-high vowels, which have a relatively wide aperture of the oral cavity and do not build up enough pressure to push the vocal folds open, are not likely to be devoiced, while high vowels, which have narrower constriction and sustain high oral air pressure, tend to be devoiced.

We also investigated the effects of the manner of articulation and the degree of aspiration of the preceding consonants. Both the manner of articulation and the degree of aspiration of the preceding consonants are important factors in vowel devoicing; fricatives tend to devoice the following vowels more often than stops or affricates and highly aspirated consonants tend to devoice the following vowels more often than others.

Frication seems to be a more important consonantal factor than aspiration in vowel devoicing. For example, among alveo-dentals, a lenis fricative, /s/--which sustains longer, narrow oral constriction--makes about 74.3% of the following high vowels devoiced, while an aspirated stop, /t/--which does not sustain a narrow oral constriction for as long as /s/--devoices 36.4% of the following high vowels. A sustained oral constriction allows the build up of sufficient oral pressure to keep the vocal folds open. An aspirated affricate which is characterized by both aspiration and some degree of frication devoices the following vowels the most frequently (88.6%).

Another interesting observation is that fortis consonants rarely devoice the following vowels (2.0%, 3/150). According to Jun et al. [5], vocal folds approximate to each other and are raised right before the oral release of fortis consonants. This early approximation of vocal folds and enlarged subglottal cavity may explain why fortis consonants rarely condition devoicing; such an explanation is consistent with a gesture overlap view of devoicing. Moreover, all three devoiced tokens in the context of a fortis C are found with a fortis fricative. This indicates that frication is indeed an important factor in vowel devoicing.

The durational measurements did not show any significant relationship between syllable duration and vowel devoicing. Although the duration of a voiced interval becomes shorter when the vowels are devoiced, as expected, the syllable durations are not shortened. This is consistent with the finding that Korean high vowels are not segmentally deleted but devoiced. If Korean high vowels were deleted, then the duration of a devoiced syllable should be significantly shorter in the devoicing (=deleting) context in comparison to other contexts. Furthermore, the duration of a voiceless interval would not necessarily be longer in the devoiced tokens.

The mean F2 of devoiced tokens for front, central, and back high vowel shows a significant difference depending on front/backness of the target vowel, which might explain the perceptual maintenance of the identity of devoiced vowels. The consistent high values of F2, however, might result from the interactions of supraglottal resonances with subglottal ones. In the subglottal formant measurement study, Cranen and Boves [10] found the second resonant frequency at around 1400 Hz from a point 2.5 cm below glottis. They also found that all the subglottal resonant frequencies increase with the increased cross-sectional area of glottis, which might be the reason of the elevated F2s of devoiced tokens.

Lastly, we observed the intervocalic voicing of /t/ which is not intervocalically positioned. Even if voicing is masked during the production of a target vowel in CVta due to the gestural overlap with the preceding consonant, the phonological gesture of it should maintain to condition the intervocalic voicing of the following consonant, /t/.

In conclusion, on the bases of all temporal, spectral observations, we claim that vowel devoicing in Korean is a phonetically driven process due to vowel undershoot by a gestural overlap. The identity of devoiced vowels is kept while vowels are devoiced, and can be recalled from the second formant information. This can be potentially utilized by native speakers of Korean in conversation.

5. REFERENCES

[1] Bloch, B. 1950. Studies in colloquial Japanese IV: Phonemics. *Language* 26, 86-125.
[2] Hirose, H. 1971. The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese. *Phonetica* 23, 156-170.
[3] Jannedy, S. 1996. Gestural phasing as an explanation for vowel devoicing in Turkish. *OSU Working papers in Linguistics* 45, 56-84.
[4] Jun, S.-A., Beckman, M.E., Lee, J. 1996. Fiberscopic evidence for the influence on vowel devoicing of the glottal configuration for Korean obstruents. *UCLA working papers* 96, 43-68.
[5] Jun, S.-A., Beckman, M.E., Niimi, S., Tiede, M. 1996. Electromyographic evidence for a gestural-overlap analysis of vowel devoicing in Korean. *UCLA working papers* 96, 1-42.
[6] Yoshioka, H. 1981. Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese. *Phonetica* 38, 236-251.
[7] Davidson, L. 2006. Schwa elision in fast speech: Segmental deletion or gestrual overlap? *Phonetica* 63, 79-112.
[8] Kondo, M. 2005. Syllable structure and its acoustic effects on vowels in devoicing environments. In: van de Weijer, J., Najo, K., Nishhara, T. (eds), *Voicing in Japanes.* Berlin, New York: Mouton de Gruyter, 229-245.
[9] Manuel, S.Y., Shattuck-Hafnagel, S., Huffman, M.K., Stevens, K.N., Carlson, R., Hunnicutt, S. 1992. Studies of vowel and consonant reduction. Proc. of *ICSLP Banff,* Canada, 943-946.
[10] Cranen, B., Boves, L. 1987. On glottal formant analysis. *JASA*, 81 (3), 734-746.

My grateful acknowledgment of insightful comments and encouragement to Prof. Jennifer Cole and Prof. Mark Hasegawa-Johnson