Acoustic Correlates of 'Devoiced' Vowels in Standard Modern Japanese

J. Kevin Varden
Meiji Gakuin University

Abstract: Loss of voicing in high vowels in Standard Modern Japanese has been traditionally characterized as a loss of voicing, more recently as a spread of [+spread glottis] from the preceding consonant. However, variability in the amount of voicing present has been observed, consistent with analyzing devoicing as gestural overlap of glottal instructions. Examination of zero crossing traces for this data set indicates gradient changes in stricture as well; fricativized voiced and devoiced vowels similar to those posited for other languages were produced by these participants. Combined with previous glottal and formant structure observations which suggest independent control of voicing and tongue shaping gestures, coordinated but independent control of oral closure as well is indicated.

LOSS OF HIGH VOWEL VOICING IN JAPANESE

The high vowels /i/ and /u/ can be devoiced in many Japanese dialects when they fall between voiceless consonants, or after a voiceless consonant at the end of an utterance (devoicing environments). They are generally held not to devoice if these conditions are not met, or if factors such as pitch accent placement interfere with the devoicing process (non-devoicing environments). Although this process was originally reported as a fast-speech rule (1), recent research has indicated that it now occurs even at slower SRS (2) (3). Vowels that have undergone loss of voicing have been characterized as 'devoiced' or 'whispered' in some contexts and 'deleted' in others. Whether or not frication is present has been used to differentiate the two cases (4). It has also been shown that the amount of voicing present is gradient (5), consistent with an analysis of devoicing as an overlap of the voicing gesture by surrounding devoicing gestures (6).

Devoicing has often been seen as a spreading of [-voice] from the preceding consonant. More recently it has been attributed to a spread of [+spread glottis] from a preceding consonant (3). It has also been noted that even when the vowel is 'deleted', formants can be observed in the frication of the preceding consonant (4).

'DEVOICED' VOWELS AS FRICATIVES

Fricativized vowels have been reported for several languages (7); the majority of consonant/devoiced vowel combinations in this data set also appear to have been made with a fricative closure. This could be characterized as spread of [+consonantal] (8), or as overlapping of stricture gestures similar to voicing gesture overlap.

If characterization of 'devoicing' as a process of fricativization due to gestural overlap is correct, rather than simple devoicing or deletion, it is predicted that gradient amounts of frication will be observed across vowel sites just as gradient amounts of voicing are. To test this, frication as indicated by zero crossings was measured at vowel sites and preceding consonants in both devoicing and non-devoicing environments.

METHODS & PROCEDURES

The collection of data used in this study is described in full elsewhere (9). Briefly, 10 young native speakers of Tokyo Japanese, ages 19-23, were recorded reading 10 sentences displayed on a computer monitor six times each at each of three controlled SRS. Each of the 10 carrier sentences contained a token with two vowel devoicing sites. For this study 1580 vowel sites were examined (22 vowel sites x 18 repetitions x 4 participants, 4 tokens not utilized) from both the target tokens and other CV material in the carrier sentences.

Frication in speech is readily observable as high-frequency noise in wide-band spectrograms. However, spectrographic displays are subject to mark level and averaging limitations and do not allow objective quantitative measurement. In addition, it was suspected that electro-palatographic sensors placed on the tongue and palate would interfere with fricative articulations. It was therefore decided to take zero crossing measurements (measures of the number of time a signal crosses the local baseline) as a measure of frication.

Zero crossing measurements (25ms window) were taken at consonant and vowel sites with the software Signalyze™. To avoid the mark level limitations of observing voice bars in wide-band spectrograms,
waveforms were used to determine vowel voicing. Signals were low-pass filtered at 500Hz and normalized to highlight weak voicing activity. When necessary to determine vowel presence, narrow-band spectrograms of the isolated, normalized vowel sites and surrounding segments were also used to observe formant structure. Zero crossing measures for various CV consonant onsets and the following vowels were taken in both devoicing and non-devoicing environments. Data was analyzed with the software package SuperAnova™.

RESULTS

The following box plot shows the general results obtained for all CVS where there was a clear distinction in the number of zero crossings at the C and V sites (1113 of each). The left side of the figure shows the number of zero crossings associated with CV mora containing a devoiced vowel (C before dvcd V, dvcd V), the right those associated with CV mora containing a voiced vowel (C before vcd V, vcd V). Each group of three boxes indicates the number of zero crossings in each voicing condition associated with each SR utilized in the data collection (f = fast; n = normal; s = slow). Outliers, the extreme 20% of the values, lie outside the boxes.

![Box Plot](image)

FIGURE 1. Zero crossings values across all segments, grouped segment type, V voicing and SR.

A three-factor ANOVA for effects of segment type, SR and vowel voicing on the number of zero crossings for these samples indicate that the interaction between type of segment and vowel voicing was significant, F(1, 2222) = 280; p < .05. The effect of SR was not significant at the .05 level. Vowel voicing was significant for both types of segments at the .05 level, F(1, 1111) = 270 for C, F(1, 1111) = 3200 for V. Generally, devoicing is accompanied by higher zero crossing levels, indicative of frication caused by a fricative closure.

Strikingly, it can be seen that the outliers cover most of the range of zero crossings—both devoiced vowels produced with low frication and voiced vowels produced with high frication can be seen. The same finding can be seen on a scattergram plot of zero crossings at vowel sites vs. fo duration (not shown); although only devoiced vowels display very high levels of zero crossings, varying levels of zero crossings are observed with vowels of all voicing durations. The same basic relationship and a range of values can be seen for each consonant and vowel of each mora checked in this study (ki, ku, shi, su, tsu) when plotted separately.

While frication and devoicing are correlated here, they are not concomitant. As with gradient amounts of voicing and presence of formants on frication preceding deleted vowels, the gradient amounts of frication observed accompanying both consonants and a following voiced or devoiced vowel indicate coordinated yet independent control of stricture as well. Further consideration of vowel devoicing as fricativization is warranted.

REFERENCES