Acoustic consequences of vowel deletion in devoicing environments

James Whang[1], Jason Shaw[2], and Shigeto Kawahara[3]
[1]Seoul National University, [2]Yale University, [3]Keio University

Background: In Tokyo Japanese, high vowels are typically devoiced between voiceless consonants (Fujimoto, 2015). Recent work tracking the movement of the tongue dorsum using Electromagnetic Articulography (EMA) has argued that the lingual gestures of vowels are also variably deleted in this context (Shaw & Kawahara, 2018b). Moreover, the probability of deletion varies systematically across words, raising the question of how such variable patterns are learned. A similar conclusion—that devoiced vowels are sometimes deleted—has been drawn on the basis of acoustic analyses (Whang, 2018). Whang (2018) showed that vowel devoicing has an influence on the spectrum Center of Gravity (CoG) of preceding consonants, possibly because vowel deletion eliminates the coarticulatory influence of the vowel on the preceding consonant. To address whether the presence/absence of a vowel can be detected in the acoustics, we applied Whang's acoustic analysis of consonants to Shaw & Kawahara's EMA data.

Data and analysis: The data came from the EMA and corresponding acoustic recordings of 6 native speakers of Tokyo Japanese, reported in Shaw & Kawahara (2018b). Each speaker produced 10-15 repetitions of five target items containing /u/ in a devoicing environment and a closely matched control with a voiced vowel, e.g., /ɸusoku/ 'shortage' ~ /ɸuzoku/ 'attachment'. The presence vs. absence of a vowel was determined for each token via Bayesian classification of tongue dorsum (TD) trajectories (Shaw & Kawahara, 2018a). The classifier was trained on voiced vowel TD trajectories (vowel present scenario) and a linear interpolation between flanking segments (vowel deletion). We measured CoG over a 20ms window centered on the point of maximum constriction for the consonant preceding the target (devoiced) vowel (see Figure 1). This measurement was chosen to replicate Whang (2018), who found higher CoG for consonants preceding devoiced vowels than preceding voiced vowel controls. Statistical differences were evaluated with linear mixed effects models fit to each consonant separately.

Results: Normalized CoG by condition is summarized in Figure 2. The effect of vowel devoicing on preceding consonant CoG reported in Whang (2018) was largely replicated in this data. Across five consonant environments /ʃ/, /s/, /ts/, /k/, /ɸ/, CoG tended to increase when followed by a devoiced vowel; however, the CoG raising effect was only significant for /ts/ and /ɸ/ contexts. The CoG for /ɸ/ was raised by an average of 1807 Hz ($p < .0001$) when the following vowel was devoiced. As for the effect of deletion on CoG, for tokens in which the lingual target of the vowel was classified as absent, the CoG of /ɸ/ increased further, by an average of 3267 Hz ($p < .0001$) relative to /ɸ/ in the voiced vowel context. However, the only other environment to show a further increase in CoG from the *devoicing-vowel present* condition to *devoicing-vowel absent* condition was /k/. The presence vs. absence of a lingual target did not affect the CoG of coronals: /ʃ/, /s/, /ts/.

Discussion: Analysis of consonant CoG indicates that it may be possible to detect the absence of a lingual target for devoiced vowels in non-coronal contexts: /k/ and /ɸ/. In these contexts, vowel deletion may actually enhance the percept of vowel devoicing by further raising CoG. For coronals, increases in CoG due to devoicing were not robust and were not enhanced by the absence of a vowel. This is possibly related to the relatively large number of coronals in Japanese. For coronals, other aspects of articulation may tradeoff with vowel devoicing/deletion to maintain spectral contrast.

Figure 1: Example of the CoG measurement point based on minimum velocity of C_1 from a token of /ɸusoku/ which was classified as lacking a vowel height target in the kinematic signal.

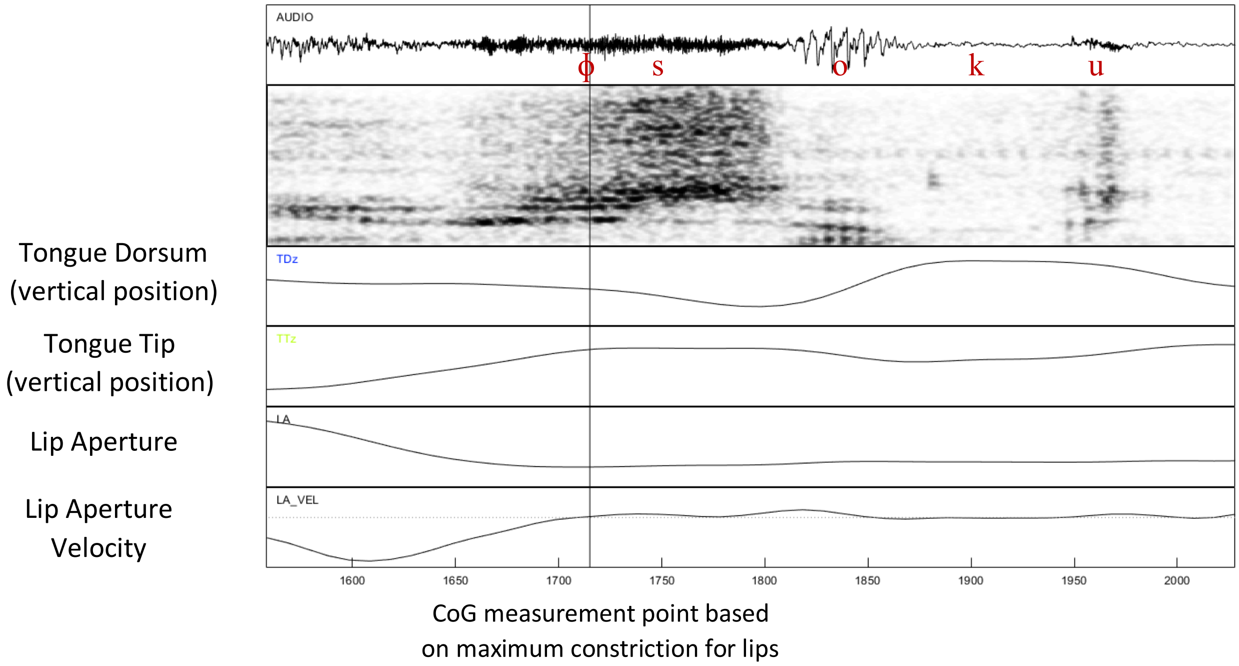
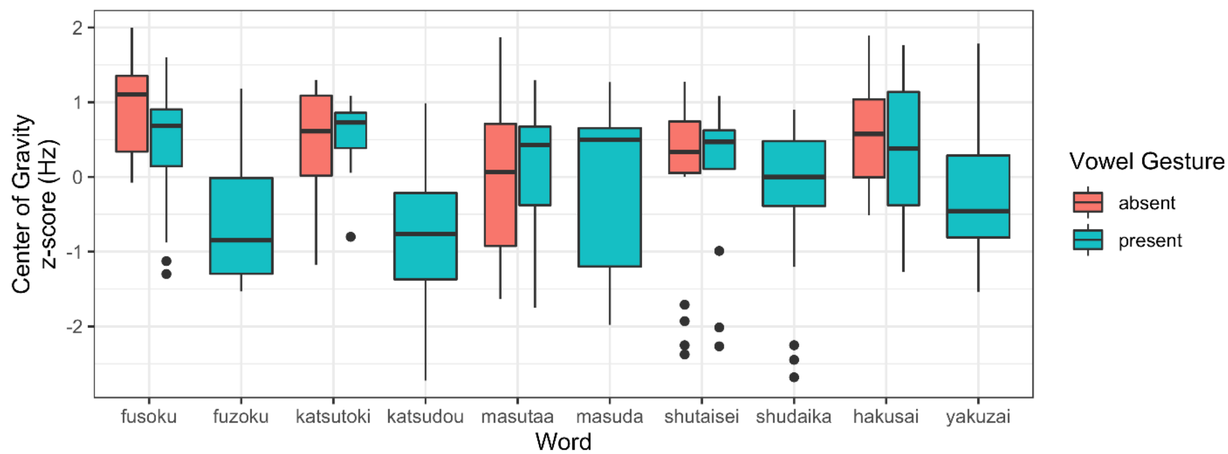


Figure 2: Normalized CoG for C_1 preceding voiced vowels and devoiced vowels with and without a vowel height target



References

Fujimoto, M. (2015). Chapter 4: Vowel devoicing. In H. Kubozono (Ed.), *The handbook of Japanese phonetics and phonology.* Berlin: Mouton de Gruyter.

Shaw, J. A., & Kawahara, S. (2018a). Assessing surface phonological specification through simulation and classification of phonetic trajectories. *Phonology, 35*(3), 481-522. doi:10.1017/S0952675718000131

Shaw, J. A., & Kawahara, S. (2018b). The lingual articulation of devoiced /u/ in Tokyo Japanese. *Journal of Phonetics, 66*, 100-119. doi:https://doi.org/10.1016/j.wocn.2017.09.007

Whang, J. (2018). Recoverability-driven coarticulation: Acoustic evidence from Japanese high vowel devoicing. *The Journal of the Acoustical Society of America, 143*(2), 1159-1172.