

Acoustic correlates to interpretation in Korean *what-so* constructions

Geonhee Lee

Stony Brook University

Background & Prediction. *What-so* constructions refer to *wh*-questions containing the expression *mwue-l kulehkey* ‘what-ACC so’, which are ambiguous between a canonical *what*-question reading (1a) and a *why*-question reading (1b) [4]. While previous studies on Korean *wh*-constructions focus on the role of IP- and AP-level tonal specification in signaling different interpretations [2,6], prosodic tones do not suffice to distinguish between the two readings in *what-so* constructions [5]. This study hypothesizes that disambiguation relies on the presence of a post-*wh* boundary. It signals the interpretation of *wh*-indeterminates: typically absent for a *wh*-pronoun (e.g., *what*) but follows an indefinite pronoun (e.g., *something*). Thus, the prediction is that the *what*-question reading lacks a post-*wh* boundary as the *wh*-word *mwues* ‘what’ is faithful to the reading, whereas the *why*-reading exhibits one, as the literal meaning of the *wh*-word does not align with the reading. To test this via domain-initial strengthening [1,3], this study examines acoustic cues at the post-*wh* boundary: closure duration (CD), voice onset time (VOT), and voicing ratio (VR) of the initial [k] in *kulehkey* ‘so’.

Data & Method. The study analyzed speech corpus data from YouGlish, a website that provides naturally occurring speech data with prosodic information and discourse context. The data were categorized into two contexts: *what*-questions, if a simple constituent answer followed, and *why*-questions, if a reason was given as the answer. A total of 23 instances (6 *what*-Qs, 17 *why*-Qs) were analyzed, all of which reliably preserved voicing information. A logistic regression model was fitted with interpretation (*what*-Q and *why*-Q) as the dependent variable, and CD, VOT, and VR as predictors. CD is indirectly associated with voicing: longer CD duration is associated with less voicing. It is also sensitive to prosodic boundaries as it is linked with the domain-initial strengthening effect [1]. VOT directly quantifies voicing: the longer the VOT, the less voiced the sound. VR measures the proportion of the closure interval during which voicing occurs, i.e., the voice duration of [k] divided by the total duration of [k].

Results. The logistic regression model revealed a marginal effect of CD on interpretation. Longer CDs were associated with a higher probability of a *why*-question reading ($\beta \approx 1.35, SE \approx 0.72, Z \approx 1.88, p \approx 0.06$), as shown in Figures 1 and 2. Despite limited statistical power, the estimated effect size was substantial (Odds Ratio = 3.86). A log-likelihood ratio also showed a significant effect of CD ($\chi^2 \approx 4.84, p \approx 0.03$). In contrast, neither VOT (Figure 2B; $\beta \approx 0.08, SE \approx 0.55, Z \approx 0.15, p \approx 0.88$) nor VR (Figure 2C; $\beta \approx 0.57, SE \approx 0.65, Z \approx 0.89, p \approx 0.38$) showed a significant effect on the interpretation. As in Figure 2, the 95% confidence intervals for these variables are wide, indicating that they do not reliably predict interpretation. A follow-up Wilcoxon rank-sum test was conducted to determine whether CD, VR, and VOT differ across interpretation types. The results confirmed a marginally significant difference in CD ($W = 23, p = 0.05$) and no significant differences in VR ($W = 47.5, p = 0.82$) or VOT ($W = 53, p = 0.92$) between the two groups.

Conclusion. This study investigated whether the interpretation of Korean *what-so* constructions is associated with phonetic correlates at the post-*wh* boundary. Among the acoustic cues, CD differed in interpretation, whereas VOT and VR did not consistently distinguish the two. These findings suggest that interpretive differences are reflected in temporal expansion (CD) rather than in laryngeal voicing properties (VOT, VR). In particular, the results are reliable with accounts in which domain-initial strengthening serves as a phonetic cue to prosodic structure [1], which, in turn, correlates with the resolution of *wh*-ambiguity.

Overall, this paper provides phonetic evidence that prosodic structuring plays a role in shaping interpretation in Korean *wh*-constructions.

- (1) Mwue-l kulehkey manhi mek-ess-e
 What-ACC so much eat-PST-SFP
 a. ‘What did (pro) eat so much of?’
 b. ‘Why did (pro) eat so much?’

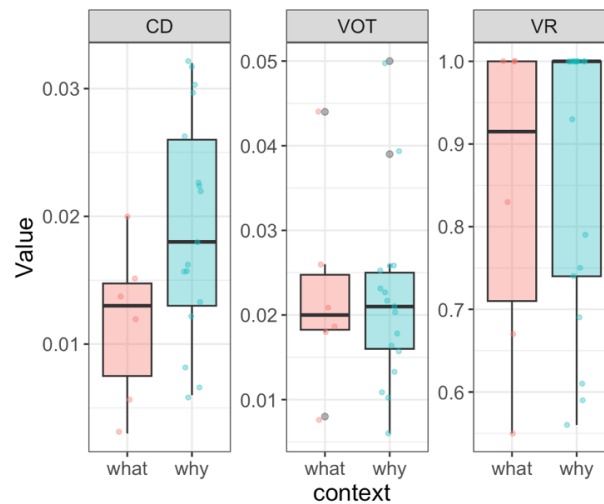


Fig. 1. The box-whisker plots of the data by interpretations.

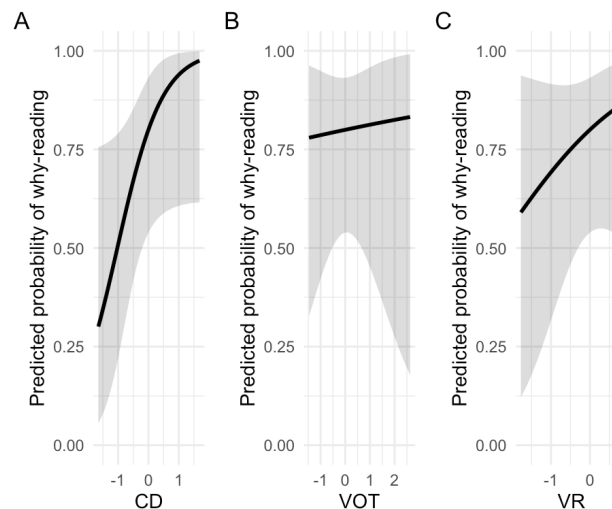


Fig. 2. Predicted probability of a *why*-question reading as a function of (A) CD, (B) VOT, and (C) VR. Solid lines represent the logistic regression fit, and shaded areas indicate 95% confidence intervals.

References

- [1] Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29, 155-190.
- [2] Jun, S-A., & Oh, M. (1996). A prosodic analysis of three types of *wh*-phrases in Korean. *Language and Speech* 39, 37-61.
- [3] Jun, S-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15, 189-226.
- [4] Kim, O. (2022). The syntax and semantics of non-standard *wh*-constructions in Korean. PhD Dissertation, University of Wisconsin, Milwaukee.
- [5] Lee, G. (submitted). Resolving the four-way ambiguity in Korean *what-so* constructions. Manuscript under review.
- [6] Yun, J. (2019). Meaning and prosody of *wh*-indeterminates in Korean. *Linguistic Inquiry* 50, 630-647.