

## A conflicting cue experiment concerning the Korean laryngeal F0-VOT contrast

Michaela Watkins<sup>1</sup>

<sup>1</sup>*University of Amsterdam*

Research on the Korean laryngeal contrast has demonstrated that aspirated stops increase in F0 whilst reducing VOT, such that VOT overlaps with lenis, with lenis then being distinguished through low F0 [e.g. 1,2,3,4]. The third type, fortis, is described as having a short VOT and high F0 (lower than aspirated), with the primary cue being short VOT. Creaky voice has been found to be categorically present in fortis stops [5], although its role in perception is unclear. Furthermore, research has shown that F0 and VOT can overlap between fortis and lenis stops [6, 7]. One may therefore question if VOT is strongest cue that resolutely distinguishes fortis from the other two. Continuing with the traditional F0-VOT continuum commonly described, a conflicting cue experiment was performed to assess the role of these cues in distinguishing the three-way laryngeal contrast. The experiment was split into free-writing followed by a forced-choice button selection task, the latter presented participants with three options: /ta, ta\*, t<sup>h</sup>a/ (lenis, fortis, aspirated), written in Korean.

**Hypotheses:** a short VOT should produce a fortis percept, although a low F0 and short VOT may induce a mixture of fortis and lenis responses due to potentially overlapping VOT [6,7]. Lenis and aspirated should be distinguished through F0. In the mid-high F0 space (~250+ Hz), there may be differences in the choice of stops for VOT (20ms – 40ms), as there may be shifts in cue boundary preferences (i.e., a conflicting space for fortis and aspirated stops).

Stimuli were created from a phrase-initial alveolar lenis token /ta/ from a native Korean female speaker with manipulated VOT and F0 (**Table 1**) in Praat [8]. Manipulated stimuli alongside unmanipulated fortis/lenis/aspirated tokens were given. Participants were recruited via *Prolific* [9], and were required to identify non-manipulated stimuli at a pass rate of +95% to be included. 29 participants were accepted. A multinomial logistical regression using the *mclogit* package [10] in *R* [11] was used with **lenis (L)** set as the reference level. Participants behaved as expected in the writing task, with 0 VOT often being written as an /a/. In the forced-choice test, increasing F0 was significant in choosing **aspirated (A)** over **L** (no interaction between VOT and F0), suggesting that F0 is the dominant cue in contrasting **A-L**. For choosing **fortis (F)** over **L**, short VOT was significant as was a higher F0, and here a significant interaction was found, suggesting an interaction of cues in making the **F** choice. In neither task did the participants strongly choose **F**, even for 0ms and 300Hz or 20ms and 275Hz, despite both being in the ‘acceptable’ **F** range [6]. Changing **A** to the reference level produced a similar result: the **F-A** distinction relies on a complex cue interaction between VOT and F0, whereas **L** and **A** are distinguished by F0. This alongside visual inspection (**Fig. 1**) suggests that the accepted perceptual space for **F** is limited when reduced to these cues, and is constrained to the absolute shortest VOT: however, it competes with **L** if the F0 is also lowered. Referring to the initial hypotheses, **A** and **L** remain distinguished by F0 alone, whereas VOT alone is not enough to create a strong fortis percept: even though some fortis tokens were selected at short VOT and high F0, lenis remained a preferred choice in this range. Using the new harmonics function in Praat, an analysis of production data confirmed that **F** stops show creak during the first 1/3 of the vowel compared to **L** and **A** as found in [5, 12], (**Fig. 2, 3**). Lenis stops are also breathier in the first 1/3. Along with the above results, two discussion points remain:

1. One can create a plausible aspirated stop from a lenis just by manipulating F0, but one cannot easily do the same for fortis by reducing VOT alone.
2. Related to the above, a distinction in voice quality found in the initial portion of the vowel may be important for category distinction, and not just short VOT.

## References and Figures

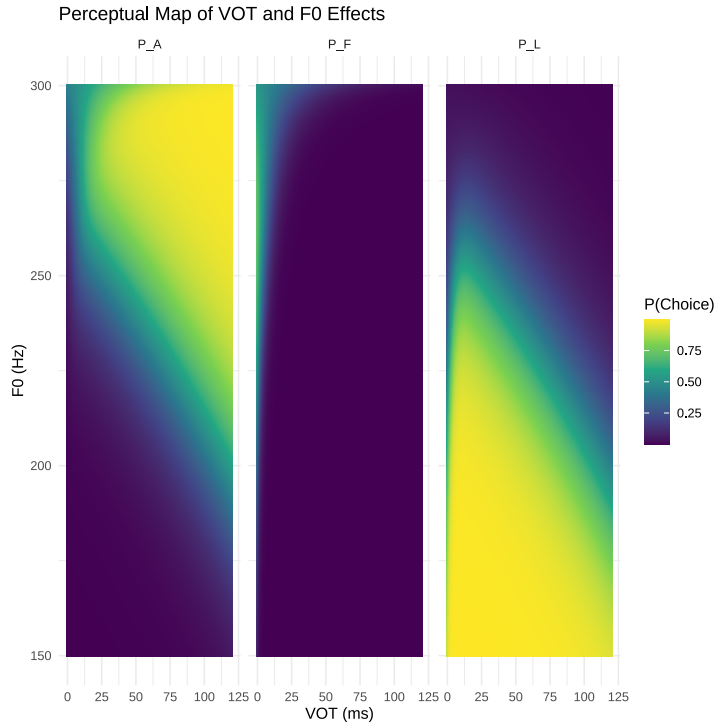


Figure 1: Perceptual map of probabilities ( $P$ ) of choosing different stop types Aspirated, Fortis, and Lenis, using VOT and F0.

|        |        |        |        |        |         |         |
|--------|--------|--------|--------|--------|---------|---------|
| F0 150 | F0 175 | F0 200 | F0 225 | F0 200 | F0 275  | F0 300  |
| VOT 0  | VOT 20 | VOT 40 | VOT 60 | VOT 80 | VOT 100 | VOT 120 |
| F0 300 | F0 275 | F0 250 | F0 225 | F0 200 | F0 175  | F0 150  |

Table 1: Stimuli for the VOT vs F0: The VOT of 60ms was found in the middle of the F0 continua, and therefore only has one manipulated stimuli. For this reason there are 13 total stimuli.

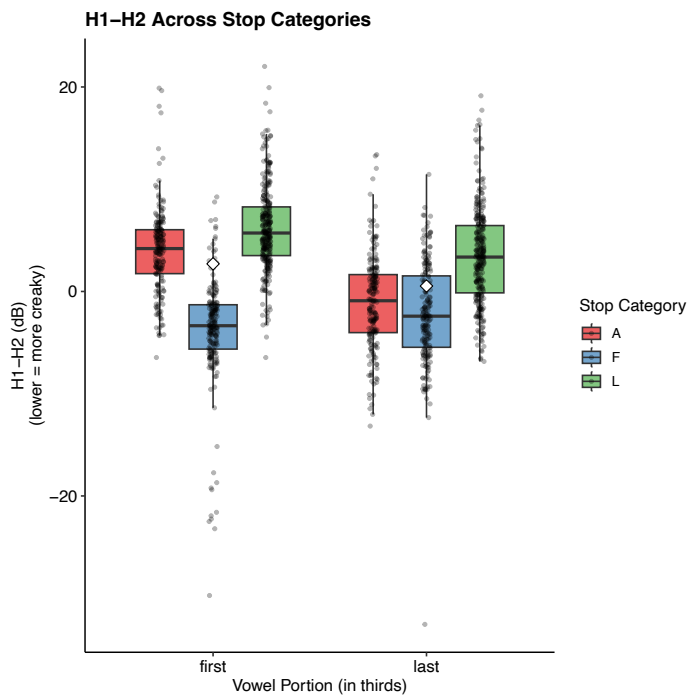


Figure 2: H1-H2 across stop categories as well as vowel position (first versus last third of vowel)

- [1] Bang, et al. (2018) The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis *J. of Phonetics* pp.120-144 [2] Silva (2006), Acoustic evidence for the emergence of tonal contrast in contemporary Korean, *Phonology* 23(2), pp.287-308 [3] Kang & Guion (2008) Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies *J. Acous. Soc. of America* [4] Kang et al. (2022) The Phonology and Phonetics of Korean Stop Laryngeal Contrasts in *The Cambridge Handbook of Korean Linguistics* pp.215-247 [5] Watkins et al. (2024) Revisiting Pitch Jumps: F0 Ratio in Seoul Korean *Proc. Interspeech 2024* pp. 3135-3139 [6] Lisker & Abramson (1965) A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements *WORD* 20(3) pp.384-422 [7] Kim, (2002) Correlation between VOT and F0 in the perception of Korean stops and affricates *ICSLP 2004* pp. 49-52 [8] Boersma & Weenink (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.47, retrieved 7 November 2025 from <https://praat.org> [9] Prolific (estab. 2014), Oct. 2025 [www.Prolific.com](http://www.Prolific.com) [10] Elff, (2026). *mclogit: Multinomial Logit Models for Categorical Responses and Discrete Choices*, R package version 0.9.17.1 [11] Posit team (2025). RStudio: Integrated Development Environment for R. Posit Software [www.posit.co/](http://www.posit.co/) [12] Watkins et al. (2025) Robustness of F0 ratio as a diagnostic: comparing creaky voice in Danish and Seoul Korean *Proc. Interspeech 2025* pp. 4203-4207

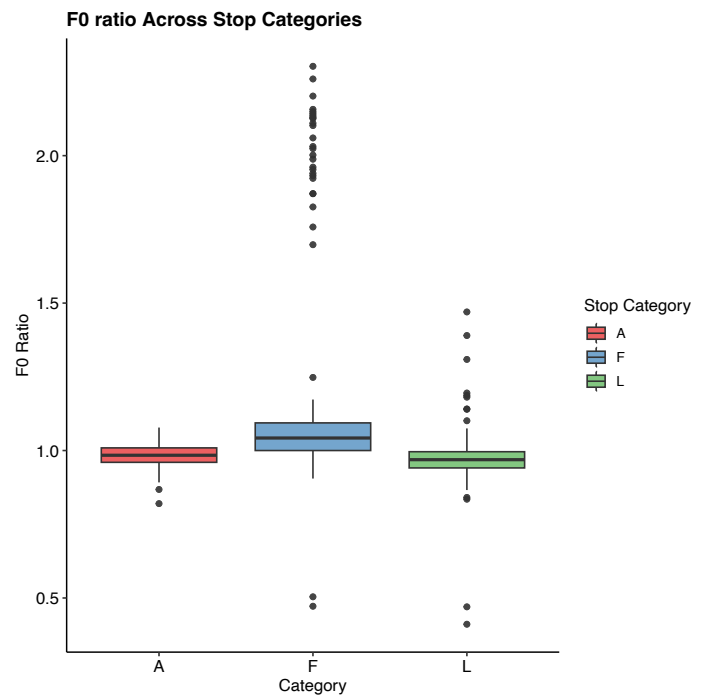


Figure 3: F0 ratio (as discussed in [5]) of the stop categories aspirated, fortis and lenis (A, F, L respectively). A value over 1.7 suggests pitch doubling / creak.