

## The Effects of VOT on Lexical Access by L1 and L2 Listeners: An Eye-Tracking Study

Eunkyung Sung<sup>1</sup>, Sunhee Lee<sup>2</sup> & Sehoon Jung<sup>3</sup>

<sup>1,2</sup>Cyber Hankuk University of Foreign Studies,

<sup>3</sup>Kyungsung University

eks@cufs.ac.kr, lishanxi@cufs.ac.kr, [schjung0427@gmail.com](mailto:schjung0427@gmail.com)

The current study examines the effects of within-category differences in voice onset time (VOT) on the dynamics of lexical activation. Specifically, this study compares sensitivity to the VOT cue between native English and Korean listeners in detecting the voicing contrast in English word-initial stop consonants. We used the eye-tracking method to monitor listeners' cognitive processes more closely when dealing with aural input during the picture identification task. The stimuli were modified natural speech tokens varying along six steps of VOT continuum for /b/-/p/, /d/-/t/, and /g/-/k/, respectively. The interval between two steps was 11–19ms for all three pairs. The participants were given aural input in the form of instructions (e.g., *look at the \_\_\_\_\_*) and asked to pick an image they just heard between the two options (i.e., *palm-bomb*, *pole-bowl*, *tart-dart*, *toe-dough*, *card-guard*, *coat-goat*) on the screen while or after they listened to the input. Figure 1 manifests the screenshot of the experiment.

The results of the experiment showed that both listener groups utilized the VOT cue in recognizing images. Figure 2 displays the results of the keyboard-click responses for each group. The response rates of voiceless stops rose as the VOT values increased for both English and Korean listeners. This means both groups utilized VOT as a major acoustic cue to identify voicing of stop categories.

Figure 3 shows the participants' proportional look on the two images — namely the target and competitor images — starting from the time they heard the target word (i.e., 560ms) and onwards. Comparing the two groups' time-course of fixations to the images, both groups showed similar profiles in that their looks to the target images including voiceless stops (i.e., /p/, /t/, /k/) increased as the VOT level increased. However, there were discrepancies between the two listener groups at least in two respects. First of all, the proportional differences between the 6 levels of VOT were much less for the Korean listeners than for the English listeners. Next, the critical point at which the Korean listeners clearly recognize the target image was found to have occurred at a later point. At the VOT level of 5 or 6, the increase of looks to target images was shown around 780 milliseconds for the Korean listeners, compared to the English listeners (about 720 milliseconds). This means that the English listeners were relatively faster than the Korean listeners in processing the target sounds. In addition, the English listeners demonstrated relatively greater and more stable focus on the targets as the VOT level increased. Therefore, despite some qualitative similarities between the two groups, the analysis of their time course data clearly revealed differences in terms of timing of image recognition and the level of certainty between L1 and L2 listeners.

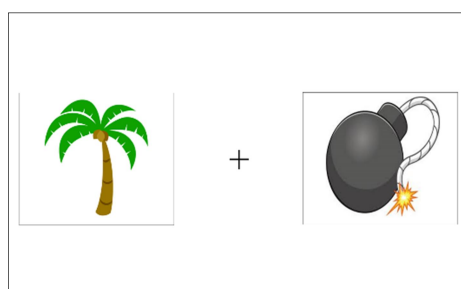


Figure 1. A screenshot of the experiment (*palm* vs. *bomb*)

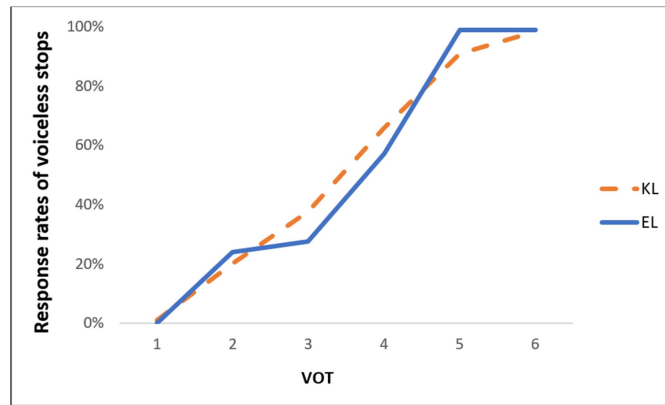


Figure 2. Response rates of voiceless stop (/p/, /t/, or /k/) as a function of VOT by English listeners (EL, solid lines) and Korean listeners (KL, dotted lines)

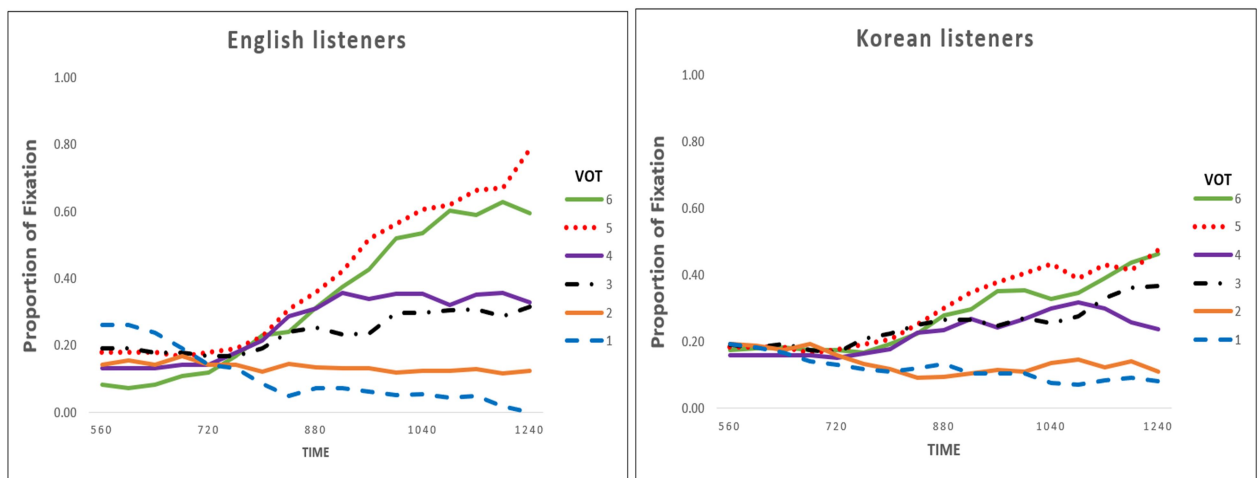


Figure 3. Mean proportion fixation to the target pictures as a function of VOT

## References

- [1] Kim S., Mitterer, H., and Cho, T. (2018). A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PLoS ONE* 13(8), e0202912.
- [2] Kong, E. J. and Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics* 59, 40-57.
- [3] Nakai, S. and Scobbie, J. M. (2016). The VOT Category Boundary in Word-initial Stops: Counter-Evidence Against Rate Normalization in English Spontaneous Speech. *Laboratory Phonology* 7(1), 13.
- [4] McMurray, B., Tanenhaus, M. K., and Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition* 86(2), B33-B42.
- [5] McMurray, B., Clayards, M. A., Tanenhaus, M. K., and Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin & Review* 15(6), 1064-1071.
- [6] Reinisch, E and Mitterer, H. (2022). Phonetics and eye-tracking. In Knight, R. A. and Setter, J. (Eds.), *The Cambridge handbook of phonetics* (pp. 457-479). Cambridge University Press.