Perception of Korean fricatives and affricates in “bubble” noise by native and non-native speakers

Korean utilizes the two-way laryngeal contrast in dental-alveolar fricatives and the three-way laryngeal contrast in affricates in terms of tenseness, a contrast that is rare in other languages (Chang, 2013). L2 Korean learners tend to have difficulty differentiating these sounds. This study uses a carefully designed type of noise to examine the cues used by L2 learners of Korean to discriminate between Korean fricatives and affricates and how these cues relate to those used by native Korean listeners.

Holliday (2012) examined the perceptual assimilation of Korean fricatives by naïve non-native listeners and novice L2 learners of Korean. The purpose of the study was to measure the perceptual similarity for these listeners between Korean and English /s/ or /ʃ/ in terms of perceptual assimilation. He also examined identification accuracy of Korean fricatives for L2 Korean learners by Minimal Pair Test and CV test. In this study, however, we aim to compare the cues used by L1 and L2 Korean speakers for perceiving tenseness.

The recent study Alwan et al. (2011) investigated speech perception for stops and fricatives in noise in order to identify acoustic cues for place of articulation. Mandel et al. (2016) conducted an experiment for speech perception in noise for American English stops perceived by native listeners of English. They created a procedure to identify the time-frequency locations in the utterance where audibility is correlated with intelligibility, which we call the “importance map” of the utterance. The methodology of Mandel et al. (2016) was chosen to test L2 speech perception of Korean fricatives and affricates because of its potential high resolution characterization of cues and relative lack of assumptions as to where these cues should be found.

In the current study, stimuli were minimal “quintet” words, which consist of consonant-vowel-consonant-vowel (CVCV) Korean words with the form of /Cada/. Specifically, the words were /sada/, /ssada/, /ʃada/, /ʃʃada/, and /ʃʃʃada/. Stimuli were selected from real speech to ensure natural production and perception. The speech tokens were mixed with speech-shaped Gaussian noise with a Signal-to-Noise Ratio (SNR) for -25dB, which is sufficient to make the speech completely imperceptible. Then, “bubbles” erasing the noise locally were created at random time-frequency locations, providing “glimpses” of the speech. The number of bubbles per second was set adaptively to achieve 60% accuracy per utterance, starting at 15 bubbles per second. Ten subjects participated, 5 native Koreans and 5 L2 Korean learners. Subjects did not have histories of hearing or visual impairment. In total, 10,000 responses were collected (10 participants × 5 consonants × 200 tokens).

The goal of the current study is to characterize the importance map of both native Korean listeners and L2 Korean learners. We hypothesize that these two importance maps will be different, indicating that the two listener types use different cues to discriminate Korean fricatives and affricates. Characterizing the cues that L2 listeners use would allow an incorrect use of cues to be diagnosed, and potentially corrected through training to perceive the differentiation of tenseness more like native Koreans. Figure 1 below shows three types of visualizations: the spectrogram of the clean utterance, the importance map overlayed on the spectrogram of the clean utterance, and the correlations themselves in areas where these correlations are significant. It shows the native Korean listener (NK5) used higher frequency regions around 4kHz to 5kHz and F1 and F2 around the vowel transition to perceive identify

---

1 During the experiment, Romanized labels for the phonemes were used in order to avoid any confusion with phonetic symbols to the participants. /s/, /ʃʃ/, /ʃʃʃ/ and /ʃʃʃʃ/ refer to, respectively, a lax (or unaspirated) fricative, a tensed fricative, a devoiced affricate, an aspirated affricate and a tensed affricate.
the word /sada/. In contrast, the L2 Korean learner (L2K1) mainly relied upon a different cue, which is frication noise around 8-9kHz. Significance of the correlation shows positive values in red which indicate that the listener identifies /sada/. For /ssada/, while the native Korean listener (NK1) used 2-4kHz frequency as an acoustic cue, frication noise around 8-9kHz and F1 and F2 on the vowel transition were used by the L2 Korean listener (L2K3). Likewise, even though the both groups tend to utilize frication noise and vowel transitions as acoustic cues to perceive Korean fricatives and affricates, these results suggest that there could be explicit differences in the use of the cues between the two groups for identifying tenseness. In addition, because L2 Korean learners have difficulty in making these discriminations, this analysis paradigm sometimes has difficulty identifying the cues that they use, although additional trials or training in this particular task would likely resolve this issue.

Figure 1. An example of the visualizations for /sada/ (NK5 for the first row and L2K1 for the second row) and for /ssada/ (NK1 for the third row and L2K3 for the fourth row)

References


