The time course of phonetic cue integration in Seoul Korean sibilant fricatives

Joohee Ko[1] and James Whang[1]

[1]*Seoul National University (Korea)*

Background: As speech unfolds over time, phonetic cues often arrive asynchronously. Previous studies have investigated the time course of how these cues are integrated for lexical access. Most of the literature supports "continuous" integration, where listeners process individual cues as soon as they arrive (e.g., [1]). Contrarily, "buffered" integration was reported for English sibilant fricatives, where processing seemed to wait for the following formant transition cue, despite the earlier frication spectra being sufficient for identifying the fricative [2]. As the Seoul Korean sibilant fricatives, nonfortis /s/ and fortis /s*/, differ in the consonantal cues (/s/ = frication followed by aspiration, /s*/ = long frication with little to no aspiration) and the vocalic cues in the following vowel (e.g., f1 transition), the current study tests the generalizability of the "buffered" integration pattern observed in English sibilant fricatives by investigating the time course of phonetic cue integration in Seoul Korean sibilant fricatives.

Methods: Thirty-one Seoul Korean listeners participated in a 4-AFC picture identification task in a visual world paradigm. Participants listened to the auditory stimuli and clicked on the corresponding image while their eye-movements were recorded. Auditory stimuli were created from four fricative minimal pairs (e.g., [sal] 'flesh' - [s*al] 'rice') produced by a female Seoul Korean speaker. For the consonantal cues, fricatives were manipulated to have fixed durations, with /s/ having 100ms of frication followed by 100ms of aspiration, while /s*/ had 200ms of frication. For the vocalic cues, rhymes produced after /s/ and /s*/ were segmented and cross-spliced onto the fricatives. This resulted in four fricative conditions – NN (/s/ consonant - /s/ vowel), NF (/s/ consonant - /s*/ vowel), FN (/s*/ consonant - /s/ vowel), FF (/s*/ consonant - /s*/ vowel). The average fixation proportions were calculated for every 4ms time bin, and $bias_{/s/-/s*/}$ was calculated by subtracting the proportion of looks to /s*/ from the proportion of looks to /s/. Values near 1 indicate more looks to /s/, while values near -1 indicate more looks to /s*/, and 0 represents an equi-biased state between /s/ and /s*/, as indicated in Figure 1.

Results: Results suggest that the integration pattern for Seoul Korean sibilant fricatives is generally continuous, with the consonantal cue effects preceding the vocalic cue effects. The cue effects were established based on the difference in $bias_{/s/-/s*/}$ between relevant conditions. Consonantal cue effects compared conditions with different consonants (e.g. NN vs. FN), while vocalic cue effects compared conditions with the same consonant but different vowels (e.g. NN vs. NF). The onset of the cue effects was established based on the time window of significant difference between conditions, obtained through a Generalized Additive Mixed Model (GAMM) analysis. The latest consonantal cue effect emerged at 535.4ms (FF vs. NF), while the earliest vocalic cue effect occurred at 565.7ms (NN vs. NF), suggesting a continuous integration pattern for Korean fricatives. To observe the effects of the different consonantal cues, another GAMM analysis was conducted for the time frame between 0 and the earliest onset of the vocalic cue effect (0-565ms) (Fig. 2). The analysis found significant nonlinear patterns over time for the conditions with the /s*/ consonant (FF, FN), indicating that the long frication led to more looks toward /s*/ before the vocalic cue effect. However, no significant pattern was observed for the conditions with the /s/ consonant (NN, NF), suggesting that fixations reflecting the effects of aspiration were equi-biased between /s/ and /s*/. Thus, results suggest that the long frication in /s*/ was more immediately used, while the aspiration in /s/ was not, with /s/ activation being delayed until the vocalic cues had an effect.

Discussion: The results suggest that Seoul Korean sibilant fricative processing is continuous, but how immediately a cue is processed may further depend on the specific cue. Moreover, a subsequent gating experiment found that participants were able to effectively use the aspiration to identify /s/ in the absence of the vocalic cues (Fig. 3), which was in line with previous offline perception studies that suggest that the aspiration cue carries more perceptual cue weight than the frication duration [3]. These results together suggest that factors specific to real-time processing may further determine how immediately a cue is processed.

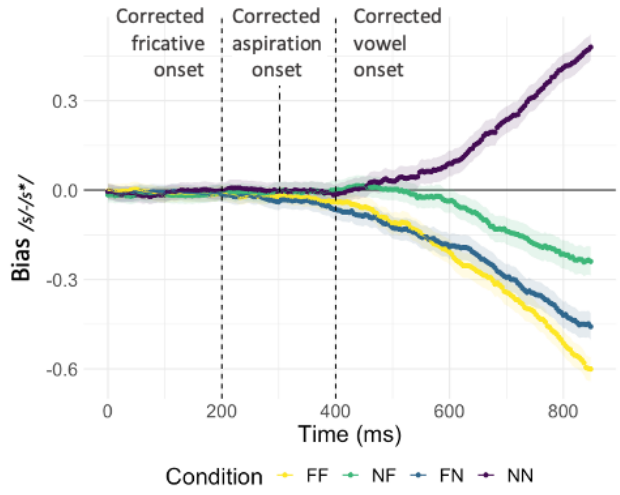

Fig. 1. Bias$_{/s/-/s*/}$ over time for the four fricative conditions. Stimuli onset is at 0ms, and the vertical dashed lines take into account the 200ms eye-movement planning, providing the corrected onsets of the fricative, aspiration, and vowel. The conditions with different consonants (e.g., FN and NN) begin to diverge before the conditions with different vowels (e.g., NN vs. NF) start to differ.

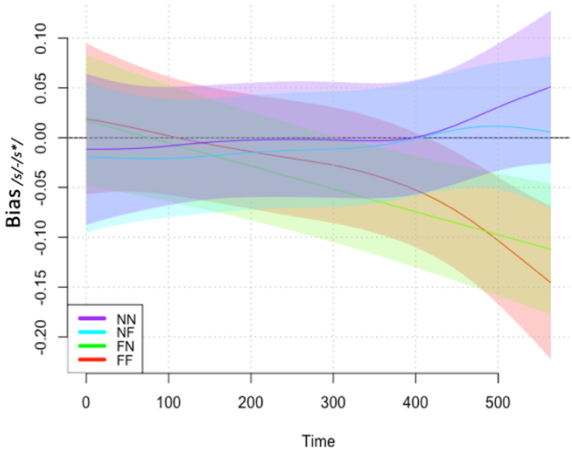

Fig. 2. Visualization of the GAMM analysis for the consonantal cue effect for the time between 0 and 565ms, before the earliest vocalic cue effect onset. While FN and FF have downward slopes, indicating more looks to /s*/ over time, NN and NF remain at 0, indicating an equi-biased state between /s/ and /s*/.

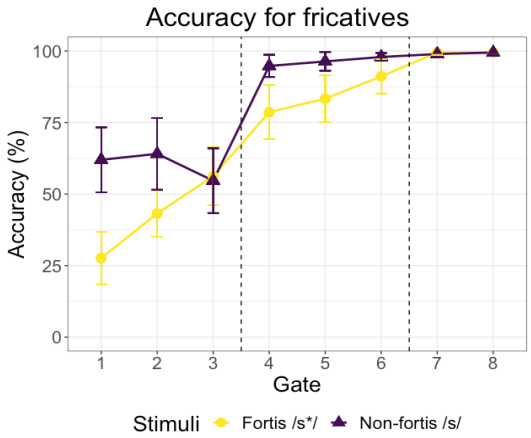

Fig. 3. Mean accuracy for each of the fricatives at each gate. The first vertical dashed line indicates the gates that begin to include aspiration for /s/ (gate 4), and the second line indicates the gates that begin to include the vocalic cues for both /s/ and /s*/ (gate 7).

References

[1] McMurray, B., Clayards, M. A., Tanenhaus, M. K., & Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic bulletin & review*, 15(6), 1064-1071.

[2] Galle, M. E., Klein-Packard, J., Schreiber, K., & McMurray, B. (2019). What are you waiting for? Real-time integration of cues for fricatives suggests encapsulated auditory memory. *Cognitive Science*, 43(1), e12700.

[3] Chang, C. B. (2013). The production and perception of coronal fricatives in Seoul Korean: The case for a fourth laryngeal category. *Korean Linguistics*, 15(1), 7–49. [https://doi.org/10.1075/kl.15.1.02cha]