White, James; Faith Chiu

Disentangling sonority and attestedness: An EEG study of onset clusters in English

Previous research shows that speakers have knowledge of the phonotactic patterns of their language, and this knowledge affects their performance in a variety of tasks (e.g., Dupoux et al., 1999). Languages differ in which sequences they allow, so phonotactics must be learned, at least to some extent, on a language-specific basis. However, researchers have also suggested that speakers are sensitive to universal principles of sonority sequencing (Berent et al. 2007). If speakers are sensitive to both language-specific phonotactic restrictions and universal sonority principles, then we might expect to see independent neural responses based on attestedness (i.e. whether a sequence exists in one's language) and sonority (i.e. marked vs. unmarked sonority profiles).

We attempted to disentangle attestedness and sonority by looking at a range of onset clusters, taken from Daland *et al.* (2011), that vary in terms of their sonority profile and attestedness in English. Here, we focus on three groups of onsets: (1) Attested, unmarked: *bl, br, kl, kr, dr, gl, gr, kw, pl, pr, tr, tw*; (2) Unattested (or marginally attested), unmarked: *bw, dw, fw, gw, thw, vl, vr, vw, pw, tl*; (3) Unattested, marked: *dg, dn, km, lm, ln, lt, ml, mr, nl, pk, rd, rg, rl, rn*. For present purposes, unmarked refers to onsets with a large sonority rise (i.e. obstruent + approximant), and marked refers to other clusters. We created two CCVC nonce words (e.g. *brip*) for each cluster, by combining the CC- cluster with two of six possible -VC rimes. The nonce words were then recorded by a phonetically trained English speaker and checked to ensure that the clusters were produced without an intervening vowel.

The experiment consisted of four parts: exposure, pre-EEG rating, passive EEG, and post-EEG rating. Participants were first exposed to all of the nonce words to give them an idea of the full range of forms that they would be hearing. These words were presented auditorily, one at a time, using insert-earphones along with orthography on a screen. Next, participants completed a nonword acceptability task, where they rated each nonce word on a scale from 1 to 8; words were again presented visually and auditorily. Previous work suggests that speakers often misperceive sequences that are illegal in their language (Dupoux *et al.* 1999). We included orthography in both the exposure phase and the first rating task to make it easier for participants to perceive the clusters properly; they were also told that all words were monosyllabic.

In the EEG recording phase, participants passively listened to the nonce words (auditory form only; no orthography) while viewing silent videos with no dialogue or subtitles. EEG recording was conducted over four blocks. Each word was presented 40 times, in a pseudo-random order, with an interstimulus interval of 1 sec. After the EEG task, participants completed a second rating task, but without orthography presented.

Rating results: Our ratings data fully replicate Daland *et al.* (2011). Participants (n=4, data collection ongoing) rated attested onsets higher than unattested onsets. Within unattested onsets, participants' ratings were strongly correlated with sonority; rising sonority clusters were rated higher than plateaus, which were rated higher than falling sonority clusters. Overall, the correlation between our ratings and those in Daland *et al.* is high, both for the pre-EEG task (r = .95) and the post-EEG task (r = .88). Furthermore, the correlation between pre- and post-EEG ratings is high (r = .93), suggesting that participants were able to detect the ill-formedness of the unattested clusters, even without orthography.

EEG results: We found divergent EEG patterns based on sonority and attestedness. In terms of sonority, we found a greater N400 (i.e. negative peak 400-500ms after onset presentation) for Unattested-Marked clusters compared to Unattested-Unmarked clusters. Attested-Unmarked clusters and Unattested-Unmarked clusters patterned similarly (Fig. 1).

The results suggest that the N400 was an event-related response to clusters that were marked in terms of sonority, but not to attestedness.

In terms of attestedness, we found a greater Late Positive Component (LPC; a positive component 600-800ms after onset presentation) for Unattested-Unmarked clusters compared to Attested-Unmarked clusters; there was no difference between Unattested-Unmarked clusters and Unattested-Marked clusters (Fig. 2). Thus, the LPC appears to be sensitive to attestedness in our study, but not sonority.

Discussion. Overall, our results indicate that sonority and attestedness are processed with some degree of independence. Specifically, clusters with marked sonority profiles resulted in an N400 compared to unmarked clusters (either attested or unattested). Moreover, unattested clusters (regardless of sonority) resulted in a greater LPC compared to attested clusters. The results suggest that listeners in our study first processed the nonce words at a phonological level that is sensitive to sonority, and later attempted to process the words at a lexical level. Since our task only contained nonce words, and participants showed an LPC to unattested *onset clusters*, this processing must involve more than just a check of the lexicon for the word itself; it must involve some notion of what is a 'possible word'.

Our results are broadly consistent with recent work looking at coda clusters in nonce words with Polish and German speakers, which (like us) found an N400 effect related to sonority profile and an LPC effect for cluster attestedness (Ulbrich & Wiese, 2015). However, Ulbrich and Wiese found an LPC for German speakers only when clusters were simultaneously marked and unattested, whereas we found an LPC for English speakers for all unattested clusters regardless of sonority. Curiously, previous studies combining real words and nonce words have found seemingly opposite results: an N400 effect associated with existence of a word and an LPC effect associated with phonotactic well-formedness (e.g., Domahs *et al.*, 2009). This disparity underscores the need for more EEG studies in this area, with real words and nonce words, to advance our understanding of the complex interaction between phonological and lexical processing.

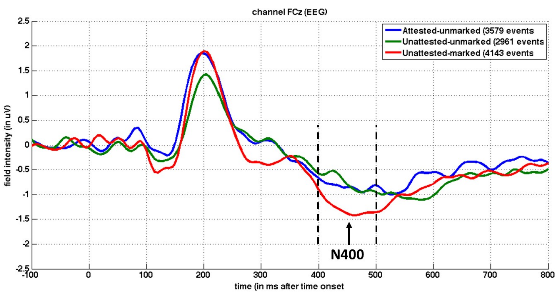

Figure 1. N400 response for marked clusters.

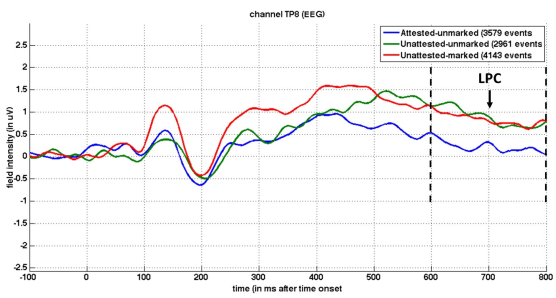

Figure 2. LPC response for unattested cluster.

References:

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition, 104,* 591–630.

Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology, 28,* 197–234.

Domahs, U., Kehrein, W., Knaus, J., Wiese, R., & Schlesewsky, M. (2009). Event-related potential reflecting the processing of phonological constraints violations. *Language & Speech, 52*(4), 415–435.

Dupoux, E. Kakehi, K., Hirose, Y., Pallier, C., & Mehler J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1568–1578.

Ulbrich, C., & Wiese, R. (2015). Processing of phonotactics constraints: A cross-linguistic EEG study. Talk presented at DGfS. Leipzig, Germany.