Overcoming the challenges of electrophysiology recordings during multi-word speech production

Svetlana Pinet[1], Nazbanou Nozari[1], Robert T. Knight[2], Stephanie K. Ries[3]

1- Johns Hopkins University, 2- University of California, Berkeley, 3- San Diego State University

Executive processes are required during language production to control accurate production and resolve potential conflicts. Sentences and word sequences in general are particularly conflicting situations, since several words must be processed and produced in a short period of time. In that context, studying sequential speech production can be challenging due to the potential mixture of processes. A response-related component that has been widely used to study conflict with EEG is the error-related negativity (ERN). Typically, its amplitude increases in higher conflict situations (e.g., for errors compared to correct trials). The aim of this study was to evaluate the feasibility of recording electrophysiological indices of monitoring such as the ERN in the context of a continuous sequence of words, and in particular, whether separate components could be isolated for each item in the sequence.

We used a tongue-twister paradigm, consisting of 32 sequences of four words sharing the same vowel and with their onset consonants arranged in an ABBA pattern (e.g., "beige tame take bale"), adapted from Oppenheim & Dell (2008). Sequences were recited twice from memory at a regular pace (a word every 650 ms). EEG data were acquired from ten participants. EEG analyses focused on correct trials, around the onset of each word. We used ICA to correct for vertical ocular movements and BSS-CCA to correct for muscular artefacts associated with speech production. Trials presenting remaining artifacts were rejected by visual inspection (20.3±9.5% trials on average), leaving 185±21 correct trials per subject on average. Laplacian transform was used as a spatial filter to enhance the resolution of components. Comparisons were performed using non-parametric statistics.

The electrophysiological data revealed a clear negative component over fronto-central electrodes, peaking 80ms after the onset of each word, similar to an ERN. Importantly, we were able to observe such distinct components for each word in the sequence. Mean amplitudes were not significantly different before and after the onset of each word ($z = -1.17$, $p = 0.24$ on FCz, between [-300:-200ms] and [+300:+400ms]), indicating a true return to baseline and a component well delineated in time. Left lateral electrodes capturing motor preparation presented a less straightforward pattern ($z = 1.89$, $p = 0.059$ on FC5), potentially suggesting more overlap between processes. An effect of repetition was evidenced over fronto-central electrodes in the [50ms:150ms] time-window ($z = 2.19$, $p = 0.028$), with a higher ERN amplitude for the second compared to the first repetition. At the behavioral level, participants made significantly more errors on the second repetition, $z = -2.7$, $p = 0.0039$.

Our results demonstrate the feasibility of isolating ERP components that index monitoring processes (ERN) at the level of single words within a multi-word speech sequence. The effect of repetition on behavioral (error rates) and electrophysiological indices (ERN amplitude) suggests that the second repetition of a sequence might be more demanding in terms of monitoring. This work paves the way to future studies of speech monitoring in contexts more complex than single word production, such as sentential production.