ERP decoding shows bilinguals represent the language of a code-switch *after* lexical processing

*Anthony Yacovone, Moshe Poliak, Harita Koya, & Jesse Snedeker (Harvard University)*
anthony_yacovone@g.harvard.edu

**Background.** For decades, research using ERPs has revealed how and when comprehenders respond to unexpected linguistic material. For example, N400 effects often occur after hearing an unexpected word, e.g., I like my coffee with cream and **salt**.[1] N400 effects can also occur after hearing an unexpected switch into another language or *code-switching*, e.g., I like my coffee with cream and **azúcar** (sugar in Spanish).[2] In a recent study, Yacovone and colleagues (2019) tested whether or not these two N400 effects are functionally distinct. To do this, they used spoken English stories with target words that varied in language (English, Spanish), contextual fit (Strong-fit, Weak-fit), or both. They reasoned that, if there were two distinct N400 effects, the weak-fitting code-switches would result in an additive effect. Results indicated that all weak-fitting conditions (regardless of language) elicited the same N400 effect. The strong-fitting code-switches, however, only elicited N400 effects in their most predictable contexts (see **Figure 1**). After initial lexical processes, all Spanish words elicited a late positive complex (LPC) and all weak-fitting words elicited a sustained negativity. Given these findings, the authors concluded three things: 1) code-switches do not elicit a *unique* N400 effect; 2) listeners can predict a particular lexical item (not just semantic features) in highly predictable contexts; and 3) bilinguals only notice that a word is in another language after the N400 time window—thus, after lexical processing.[1]

**A tempting conclusion.** This study demonstrates that the N400 is not sensitive to the language of the unexpected words per se. The only component that differentiated the language of the words was the LPC, which fully emerged *after* the N400. A tempting conclusion from these findings would be that bilinguals do not initially represent the language of words (or detect the language switch) in the earliest stages of lexical processing. However, there are two issues with this conclusion: First, traditional ERP methods cannot disentangle overlapping components; thus, it is possible that early ERP signals of the language switch were present but simply overwhelmed by the robust N400 effects. Second, ERPs cannot tell us anything about the type of information being represented in an individual's neural signals. The core issues are that 1) overlapping components cannot be disentangled, and 2) that the sensitivity of an ERP component does not reveal what information is or is not being processed at any given moment. In order to answer such questions, we would need to use neural decoding techniques.

**The present study.** We used *information-based* decoding[5,6] to assess if (and when) information about a word's language is present in each bilingual's neural signal. To do this, we decoded each participants' data separately at each time point between -200 to 2000ms. First, we collapsed all conditions into groups of English and Spanish words. Then, we split a participant's data into a training set and a testing set. The training set was then fed to a support vector machine (SVM) classifier, which used 3-fold cross-validation to create a model of the data. This model was used to predict the language of the words in the testing set given the ERP data. We recorded the accuracy of the model's predictions at each time point (20ms intervals). After analyzing all of the participants' data, we averaged together the decoding accuracies. Finally, we tested the performance of this decoding procedure against chance using a cluster-mass permutation test. Decoding accuracy was significantly above chance at distinguishing between English and Spanish words from 740-1600ms ($t = 126.55$, $p < .001$; see **Figure 2**). This cluster coincides with the LPC effect, which occurred between 750-2000ms in the original study. These results show that information about a word's language is not represented in bilinguals' neural signals until *after* lexical processing. One potential limitation is that not *all* cognitive processes are captured by EEG, leaving the possibility that language representations are present but not observable in the EEG data. Our findings have many implications for bilingualism and language processing research: We show that a bilingual is not someone with two separate and competing languages living in their mind. Rather, bilinguals are individuals with a language system optimized to handle two coding systems, where a single lexical concept can be readily mapped onto two distinct forms.

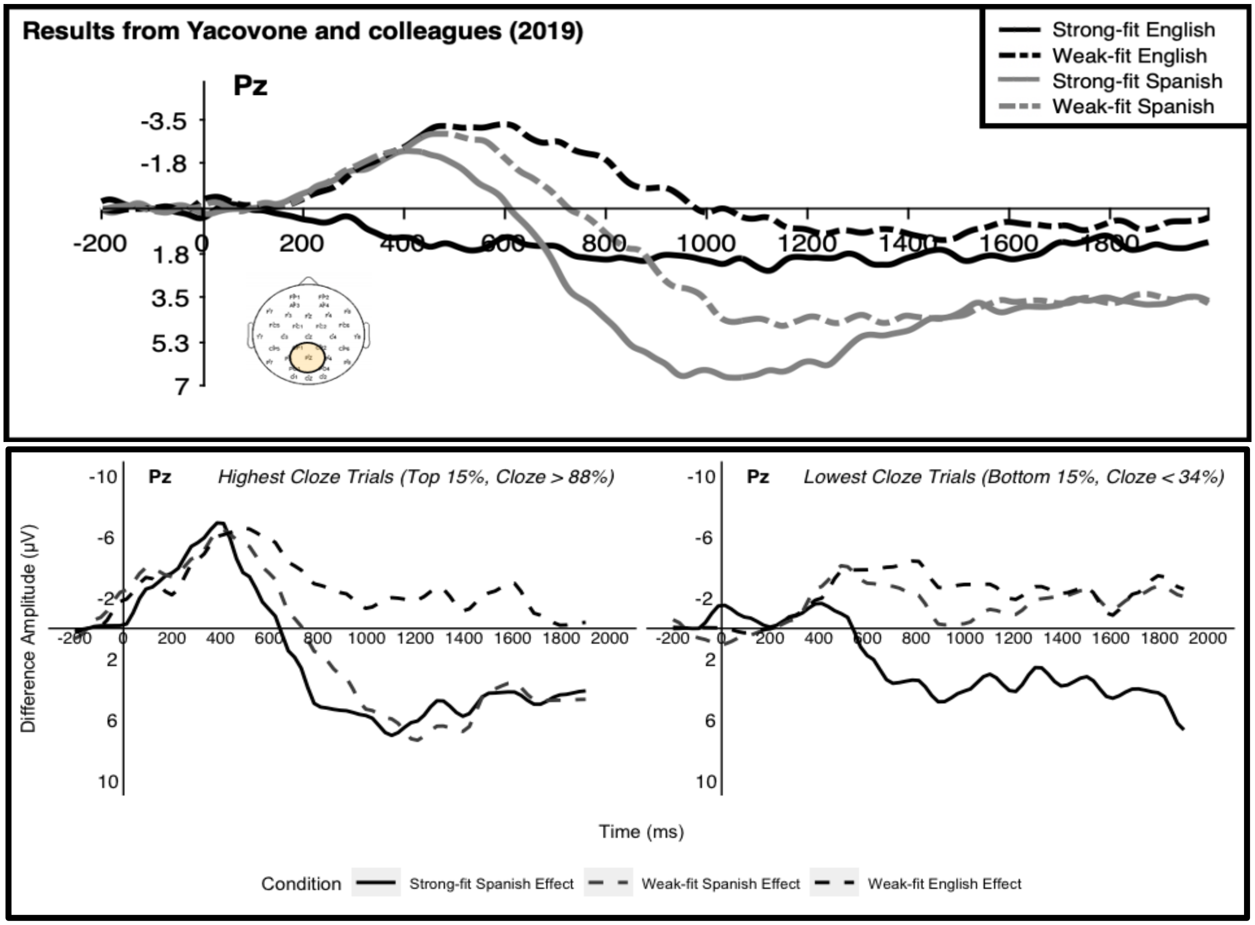


**Figure 1:** (Top panel) *Grand Averages from Yacovone and colleagues (2019)*—Spanish-English bilinguals heard sentences like "And the wig was so hot and heavy on my **head** / cranium / cabeza / cráneo." All violations resulted in the same effect during the N400 time window (250-500ms) when collapsing across all trials. All Spanish conditions elicited LPCs (750-2000ms) and all weak-fitting words elicited sustained negativities (550-1300ms). (Bottom panel) When split by cloze probabilities, the N400 effect for the Strong-fitting Spanish condition was prominent in highest cloze trials and absent in the lowest cloze trials. Thus, bilinguals predicted a specific lexical item (the strong-fitting English word) in highly predictable sentences and only semantic features (or nothing) in lowest ones.

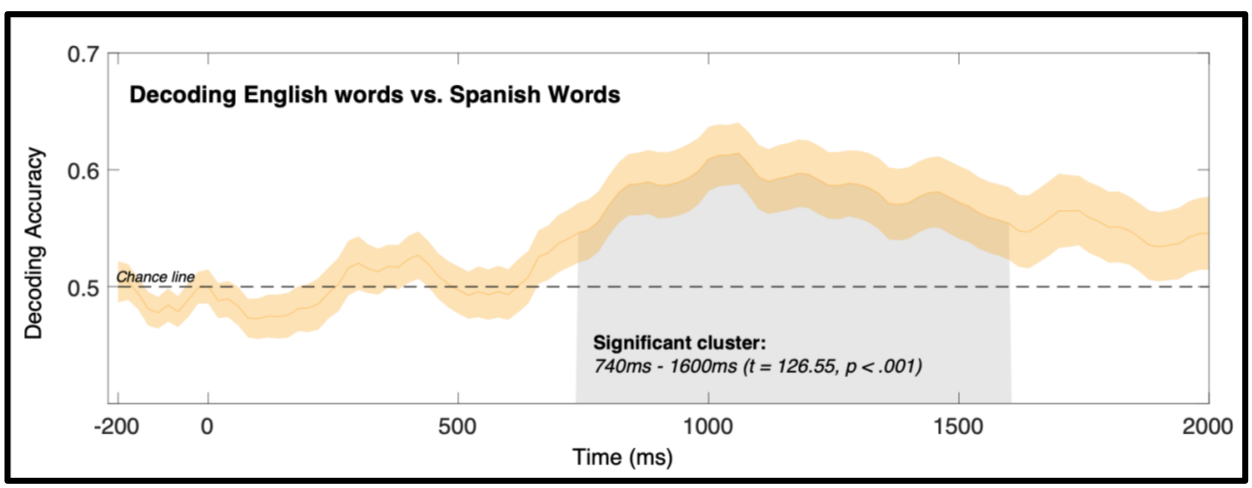


**Figure 2:** *Results from the information-based decoding technique.* This graph represents the grand average of the decoding accuracies from all 30 participants. Decoding accuracy is at chance for the first 740ms. The cluster in gray represents when decoding accuracy was significantly above chance as indicated by a cluster-mass permutation test with 1000 iterations. This timing of this cluster nicely corresponds to the emergence of the LPC effect, which was the only ERP signature that distinguished between English and Spanish stimuli in the original study.

**References:** [1]Kutas & Federmeier, 2011; [2]Litcofsky & van Hell, 2017; [3]Yacovone, Moya, & Snedeker, 2019; [4]Connolly & Phillips, 1994; [5]Luck, Bae, & Simmons, 2020; [6]Bae & Luck, 2019.