Local context sharpens the neural representation of syntactic expectancies but not of structural integration

Cosimo Iaia (iaia@psych.uni-frankfurt.de)[1,2], Alessandro Tavano[1,2]

[1]Department of Psychology, Goethe University Frankfurt, [2]Cooperative Brain Imaging Center, Frankfurt am Main

Language processing is widely understood as incremental, yet how words are integrated online into an evolving memory-based representation is unclear. Two major operations have been proposed: prediction and integration [1]. These two processes have been formalized within the framework of Surprisal Theory (for predictions/expectations) [2] and Dependency Locality Theory (for memory/integration) [3]. To dissociate their neural correlates, we used corpus-derived transition probabilities to model the online effects of prior knowledge of abstract syntactic structures and language statistics in a neural decoding analysis of an MEG dataset of participants listening to audiobooks.

Methods

We analyzed 25 participants from an openly accessible MEG datasets [4]. First, we extracted 7 measures of syntactic features commonly used in neuroscience and computational linguistics, grouped into memory features (tree depth, last right, open nodes, dependency depth) and integration features (closing nodes, dependency distance, absolute dependency). Second, to incorporate the facilitatory effects of predictions, we modelled prior knowledge by extracting marginal transition probabilities of these features at 5 different context points from a larger corpus [5]. Third, we fit a decoding model (Ridge Regression) to MEG data and compare the decoding scores of these features. Finally, we analyzed these scores using Generalized Additive Mixed Models (GAMMs) with an Autocorrelation Structure (AR1).

Results and discussion

Prior knowledge increases the decoding scores for memory related features but not for integration features (see Fig.1). However, such effect is localized to the immediate context (previous two syntactic items), and decreases for larger context windows, suggesting a sharpening of the underlying syntactic structures kept open for future completion. Integration measures were instead best decoded as word-level operation, showin no prior knowledge benefit. Such dissociation is consistent regardless of whether neural decoding is centered at word onset or offset. In line with recent theoretical proposals [6], our results provide neural evidence for a decay function in using contextual syntactic information for comprehension.

References

[1] DeLong KA, Urbach TP, Kutas M. Probabilistic word pre-activation during language comprehension685 inferred from electrical brain activity. Nature Neuroscience. 2005;8(8):1117–1121. https://doi.org/68610.1038/nn1504.

[2] Roger Levy**.** Expectation-based syntactic comprehension. *Cognition*, 2008.

[3] Edward Gibson**.** Linguistic complexity: Locality of syntactic dependencies.*Cognition*,1998.

[4] Laura Gwilliams, Graham Flick, Alec Marantz, Liina Pylkkänen, David Poeppel, and Jean-Remi King**.** Introducing MEG-MASC a high-quality magneto-encephalography dataset for evaluating natural speech processing. *Scientific Data*, 2023.

[5] W. Nelson Francis and Henry Kucera**.** A standard corpus of present-day edited American English, for use with digital computers. *Brown University, Providence*, 1964.

[6] Richard Futrell, Edward Gibson, and Roger P. Levy**.** Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science*, 2020.

Figures and Tables

Figure 1: Context averages of decoding scores for memory and integration measures at onset and offset. Base decoding indicates the raw syntactic features associated with the current syntactic state of a word, while different context (e.g., -1, -2) are the decoding scores of marginalized transition probabilities. Decoding is evaluated as the Pearson correlation for predicted features and true features and then normalized (minmax) for the whole figure.