

Context-based learning of novel meanings after minimal exposure to natural text: an EEG study

Giulia Loca¹, Fabio Marson², Marco Ciapparelli¹, Marco Marelli¹

¹ Department of Psychology, University of Milano-Bicocca, Milan, Italy

² CIMEC – Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

Background: In everyday life, we can use language both to communicate and to express novel meanings, expanding our capacity to process semantic content through linguistic information. In this scenario, language serves as a source of experience to learn new concepts that we have never encountered before through other modalities (e.g., sensorimotor; [1]). Distributional properties of language have been found to support novel meaning integration during sentence processing (e.g., via sentential context [2]). Data-driven computational measures of such regularities (e.g., surprisal or distributional similarity obtained by language models) can predict event-related potential amplitude elicited by word processing in context [3, 4]. Therefore, in this study we investigate whether minimal exposure to naturalistic text can foster the integration of novel meanings observable in the modulation of N400 amplitude, and if data-driven computational measures can predict such modulation.

Method: We used textual stimuli known as “linguistic chimeras”, i.e., language-based representations constructed by merging two existing concepts [2]. For example, we could replace the word cat in “the cat was chasing the mouse” and eagle in “the sky is crowded by eagles flying” with a single novel pseudoword (e.g., mohalk) to produce a chimera, which would probably be represented as an animal that eats mice and can fly. We identified a total of 54 different concepts from a feature-based norms database [5], paired in 27 chimeras, and extracted naturalistic sentences related to each concept from the ItWac corpus [6]. 45 participants read sentences on a screen in rapid serial visual presentation in a total of 189 trials. In each trial, two sentences (one for each concept of the chimera) provided the semantic context of the chimera (encoding phase). A third sentence, that could be either congruent (semantically related/same label), incongruent (semantically unrelated/same label), or control (semantically unrelated/different label), was shown after (testing phase). The presence of a congruency effect (i.e., stronger N400 following the novel label in incongruent condition compared to congruent ones) would signal the success of semantic integration. Additionally, three data-driven computational measures were adopted to predict ERPs amplitude: two data-driven measures of cosine similarity (one from word2vec and one from Italian BERT) and surprisal (from Minerva, an Italian autoregressive LLM). Congruency was assessed using a cluster-based permutation analysis approach, while the contribution of computational measures in the temporal profile of the N400 was assessed by model comparison.

Results: A significant Condition (incongruent > congruent) was observed at a topography and latency compatible with the N400 (Figure 1). At a more fine-grained level, Condition and BERT cosine similarity predicted amplitude in all the time windows related to N400 (200~400 ms), with BERT outperforming Condition in the earliest time window (200~250 ms) (Figure 2). Surprisal was also found to predict amplitude in earlier time-windows (200~350 ms).

Discussion: Our results demonstrate that minimal exposure to naturalistic written sentences promotes the integration of novel meanings, a process that is observable at the neural level in a time-window and topography compatible with the N400. Computational measures suggest that language distributional properties (when accounting for sentential context, i.e., BERT & Surprisal > word2vec) play a role in guiding novel meanings processing, especially in the early stages of novel words retrieval. Taken together, these results fit with the idea that the semantic system exploits linguistic experience to facilitate the learning process of novel word meaning, even after minimal exposure.

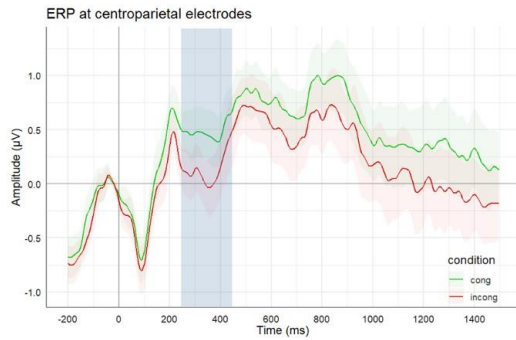


Figure 1: Event-related potential time course related to the presentation of the novel words in the testing phase.

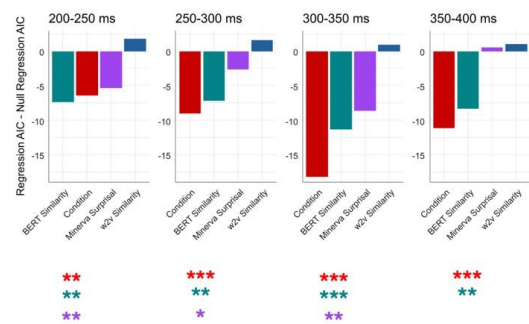


Figure 2: Results of the model comparison analysis in terms of AIC.

References

- [1] Dove, G. (2023). Rethinking the role of language in embodied cognition. *Philosophical Transactions of the Royal Society B*, 378, 20210375. <https://doi.org/10.1098/rstb.2021.0375>
- [2] Lazaridou, A., Marelli, M., & Baroni, M. (2017). Multimodal Word Meaning Induction From Minimal Exposure to Natural Text. *Cognitive Science*, 41(S4), 677–705. <https://doi.org/10.1111/cogs.12481>
- [3] Frank, S., & Willems, R. (2017). Word predictability and semantic similarity show distinct patterns of brain activity during language comprehension. *Language, Cognition and Neuroscience*, 32(9), 1192–1203. <https://doi.org/10.1080/23273798.2017.1323109>
- [4] Michaelov, J., Bardolph, M., Petten, C. V., Bergen, B., & Coulson, S. (2024). Strong prediction: Language model surprisal explains multiple N400 effects. *Neurobiology of language*, 5(1), 107–135. https://doi.org/10.1162/nol_a_00105
- [5] Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2013). Semantic memory: A feature-based analysis and new norms for Italian. *Behavior Research Methods*, 45(2), 440–461. <https://doi.org/10.3758/s13428-012-0263-4>
- [6] Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. Language resources and evaluation.