
How to talk to ChatGPT: Methodology for probing new word production in large language models

Alex Stasica & Bruna Barbosa Louzada

Utrecht University

a.stasica@students.uu.nl, b.barbosalouzada@students.uu.nl

In language, humans create words to name new objects and redefine concepts. This word creation process is a challenge in NLP in which models must capture human behaviour. Nowadays, large language models (LLMs) acquire morphological (Lane Schwartz and Tyers, 2022) and semantics (Søgaard, 2022) knowledge, and become very good at in-context prediction of words (Piantadosi, 2023) thanks to training and memorization based on large quantities of data (Piantadosi, 2023).

With the recent release of ChatGPT 3.5 (OpenAI, 2021), these models have raised questions about their ability to capture human behaviour in neologism creation. Understanding the linguistic abilities of LLMs can provide valuable insights into human behaviour from a cognitive perspective. The act of coining new words involves various mechanisms and relies on the ability to adapt to contextual and shared understanding factors. LLMs, lacking such extra linguistic information, can serve as an evaluation of the importance of this information in word coining.

Our study examined word generation by ChatGPT 3.5 and humans using two systematic prompting methodologies. The aim was to investigate if prompts providing contextual completeness improved word creation accuracy. Results indicate that context did not significantly enhance accuracy but did help generate previously rejected words. When comparing accuracies, we observed that the model performed similarly to humans, despite requiring being prompted twice for the same level of accuracy without context. These findings suggest that the LLM's behaviour resembles that of a human, even if further research is needed to understand its internal functioning.

References: • Haley, C. Schwartz, C & Tyers, F. (2022), *How to encode arbitrarily complex morphology in word embeddings, no corpus needed*. pages 64– 419 76. • OpenAI (2021), *Chatgpt (gpt-3-5)* • Piantadosi, S. (2023), *Modern language models refute Chosmky's approach to language* • Søgaard, A. (2022), *Understanding models understanding language*