

Large Language Models/Generative AI: Which/Who is This Agent?

Our aim is to show incompatibilities in the use of the term generative. The question is: What does Generative AI generate that makes it an intelligent agent? Our hypothesis is the generative character of Generative AIs is not equivalent to the generative character of human language, which creates significant differences between them on a linguistic level. The theoretical basis adopted is the Minimalist Program of Generative Grammar (CHOMSKY, 1995).

It is important to distinguish the term “generative” used to refer to the ChatGPT system from the same term used for human language, as Chomsky (1957) conjectured. LLM is a way of generating texts and extracting information from a large database using Natural Language Processing (NLP). The human linguistic system can generate an infinite number of sentences from a finite number of linguistic elements.

The data analysis is a comparison of grammaticality judgments in sentences in Brazilian Portuguese and English in ChatGPT, which are limited to pronominal coreferences and structural ambiguity. In the first experiment, with the English sentence “The child looked at the lady using the magnifying glass”, the chat identified the ambiguity and said that both “the child” and “the lady” could be using “the magnifying glass”. In Portuguese, although both the colleague and I could be exhausted, the chat replied that it was only the colleague who was exhausted in “ajudei a colega exausta no final do dia” (“I helped the exhausted colleague at the end of the day”).

In experiment 2, we tested pronominal coreferences. In English, with the sentences “Paul has said that he will travel” and “Paul has said that he will travel”, the chat identified both the coreference in the first sentence and the ungrammaticality of the second. In Portuguese, the chat both identified the coreference and correctly judged the sentences “Paulo disse que ele vai viajar” and “Paulo disse que vai viajar” as grammatical, even though there was a *pro* in place of the filled pronoun, which occurs in *pro-drop* languages.

The chat fails in the Linguistic explanation. In experiment 2, the explanation was wrongly attributed to a violation of verb tense, when the point is the null subject parameter of English and Portuguese.

The terms “agent”, “language” and “generative” are inappropriately used to refer to the Open AI tool, having no anchorage in the theoretical framework of linguistics, as Chomsky (2023).

Keywords: Generative AI X Generative Language; Language and Processing; ChatGPT; Natural and Artificial Language; Generative Linguistics and Artificial Tools

References

Chomsky N. The Minimalist Program. The MIT Press, Cambridge, 1995.

Chomsky N. Syntactic Structures. The Hague: Mouton, 1957.

Chomsky N, Roberts Ian, Watumull J. The False Promise of ChatGPT. The New York Times, March 8, 2023.

Available at: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>