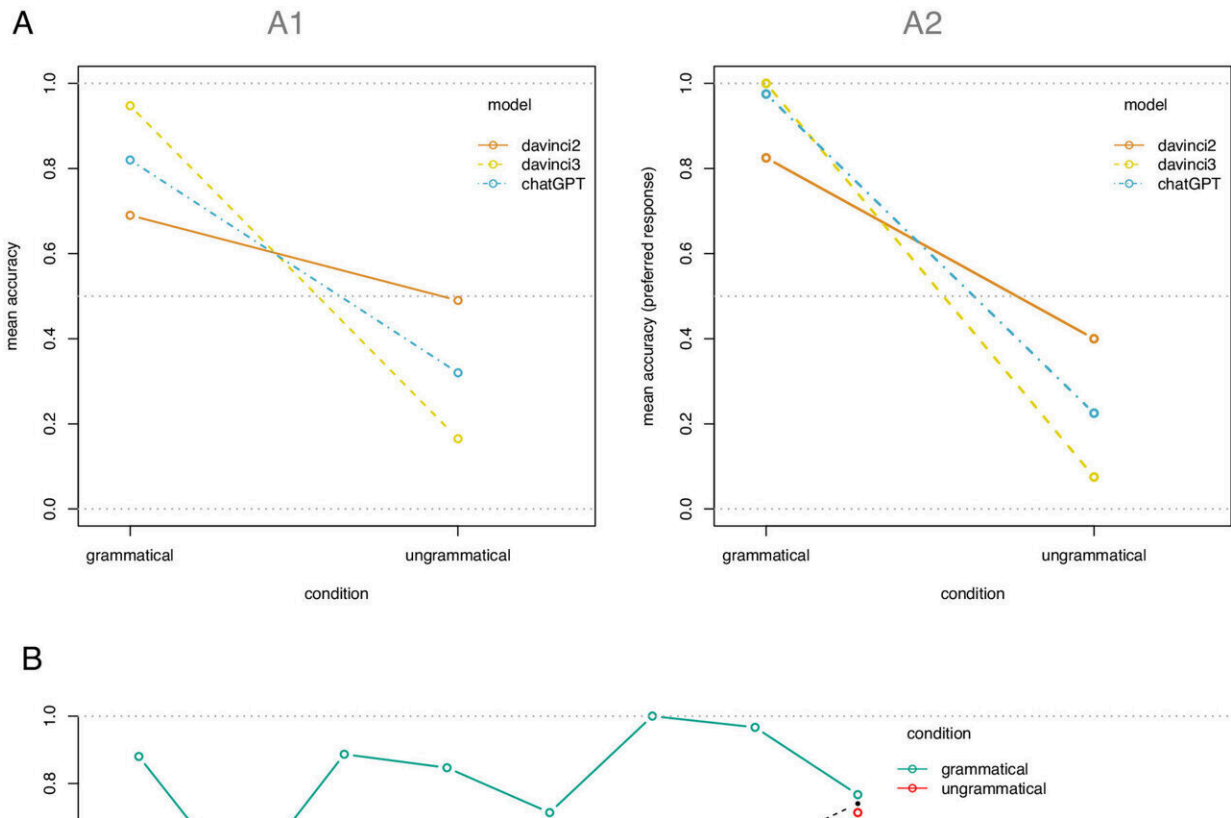


Research shows artificial intelligence fails in grammar

January 11 2024



(A) Mean accuracy by condition and model: (A1) individual responses; (A2) preferred responses per sentence. (B) Mean accuracy by phenomenon and condition. The dashed black line indicates the mean accuracy for each phenomenon across both conditions. Credit: *Proceedings of the National Academy of Sciences* (2023). DOI: 10.1073/pnas.2309583120

A study of UAB and URV researchers and [published in PNAS](#) shows that human beings recognize grammatical errors in a sentence while AI does not. Researchers have compared the skills of humans and the three best large language models currently available.

Language is one of the main features that differentiates human beings from other species. Where it comes from, how it is learned and why people have been able to develop such a complex communication system has raised many questions for linguists and researchers from a wide variety of research fields.

In recent years, considerable progress has been made in trying to teach computers language, and this has led to the emergence of so-called large language models, technologies trained with huge amounts of data that are the basis of some artificial intelligence (AI) applications: for example, search engines, machine translators or audio-to-text converters.

But what language skills do these models have? Can they be compared to those of a human being? A research team led by the URV with the participation of Humboldt-Universitat de Berlin, the Universitat Autònoma de Barcelona (UAB) and the Catalan Institute of Research and Advanced Studies (ICREA) tested these systems to check if their language skills can be compared to those of people. To do so, they compared the skills of humans and the three best large language models currently available: two based on GPT3, and one (ChatGPT) based on GP3.5.

They were given a task that was straightforward for people: they were asked to identify on the spot whether a wide variety of sentences were grammatically well-formed in their [native language](#). Both the humans who participated in this experiment and the language models were asked a very simple question: "Is this sentence grammatically correct?"

The results showed that humans answered correctly, while the large language models gave many wrong answers. In fact, they were found to adopt a default strategy of answering "yes" most of the time, regardless of whether the answer was correct or not.

"The result is surprising, since these systems are trained on the basis of what is grammatically correct or not in a language," explains Vittoria Dentella, researcher from the Department of English and German Studies, who led the study. Human evaluators train these large language models explicitly about the grammaticality of the constructions they may encounter.

By means of a learning process reinforced by human feedback, these models are given examples of sentences that are not grammatically well constructed and then given the correct version. This type of instruction is a fundamental part of their "training." On the other hand, this is not the case in humans. "Although the people who bring up a baby may occasionally correct how it speaks, they do not do so constantly in any language community the world over," she says.

Therefore, the study reveals that there is a double mismatch between humans and AI. People do not have access to "negative evidence"—about what is not grammatically correct in the language being spoken—whereas large language models, through human feedback, do. But, even so, the models cannot recognize trivial grammatical errors, whereas humans can instantly and effortlessly.

"Developing useful and safe artificial intelligence tools can be very helpful, but we need to be aware of their shortcomings. Since most AI applications depend on understanding commands given in [natural language](#), determining their limited understanding of grammar, as we have done in this study, is of vital importance," notes Evelina Leivada, ICREA Research Professor from the UAB's Department of Catalan

Studies.

"These results suggest that we need to critically reflect on whether AIs really have [language skills](#) similar to those of people," concludes Dentella, who considers that adopting these language models as theories of [human language](#) is not justified in their current stage of development.

More information: Vittoria Dentella et al, Systematic testing of three Language Models reveals low language accuracy, absence of response stability, and a yes-response bias, *Proceedings of the National Academy of Sciences* (2023). [DOI: 10.1073/pnas.2309583120](https://doi.org/10.1073/pnas.2309583120)

Provided by Autonomous University of Barcelona

Citation: Research shows artificial intelligence fails in grammar (2024, January 11) retrieved 13 May 2024 from <https://techxplore.com/news/2024-01-artificial-intelligence-grammar.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--