Cognitive scientists develop new model explaining difficulty in language comprehension

December 22 2022

Credit: AI-generated image ([disclaimer](#))

Cognitive scientists have long sought to understand what makes some sentences more difficult to comprehend than others. Any account of language comprehension, researchers believe, would benefit from understanding difficulties in comprehension.

In recent years researchers successfully developed two models explaining two significant types of difficulty in understanding and producing sentences. While these models successfully predict specific patterns of comprehension difficulties, their predictions are limited and don't fully match results from behavioral experiments. Moreover, until recently researchers couldn't integrate these two models into a coherent account.

A new study led by researchers from MIT's Department of Brain and Cognitive Sciences (BCS) now provides such a unified account for difficulties in language comprehension. Building on recent advances in machine learning, the researchers developed a model that better predicts the ease, or lack thereof, with which individuals produce and comprehend sentences. They recently published their findings in the *Proceedings of the National Academy of Sciences*.

The senior authors of the paper are BCS professors Roger Levy and Edward (Ted) Gibson. The lead author is Levy and Gibson's former visiting student, Michael Hahn, now a professor at Saarland University. The second author is Richard Futrell, another former student of Levy and Gibson who is now a professor at the University of California at Irvine.

"This is not only a scaled-up version of the existing accounts for comprehension difficulties," says Gibson; "we offer a new underlying theoretical approach that allows for better predictions."

The researchers built on the two existing models to create a unified theoretical account of comprehension difficulty. Each of these older models identifies a distinct culprit for frustrated comprehension: difficulty in expectation and difficulty in memory retrieval. We experience difficulty in expectation when a sentence doesn't easily allow us to anticipate its upcoming words. We experience difficulty in memory

retrieval when we have a hard time tracking a sentence featuring a [complex structure](#) of embedded clauses, such as: "The fact that the doctor who the lawyer distrusted annoyed the patient was surprising."
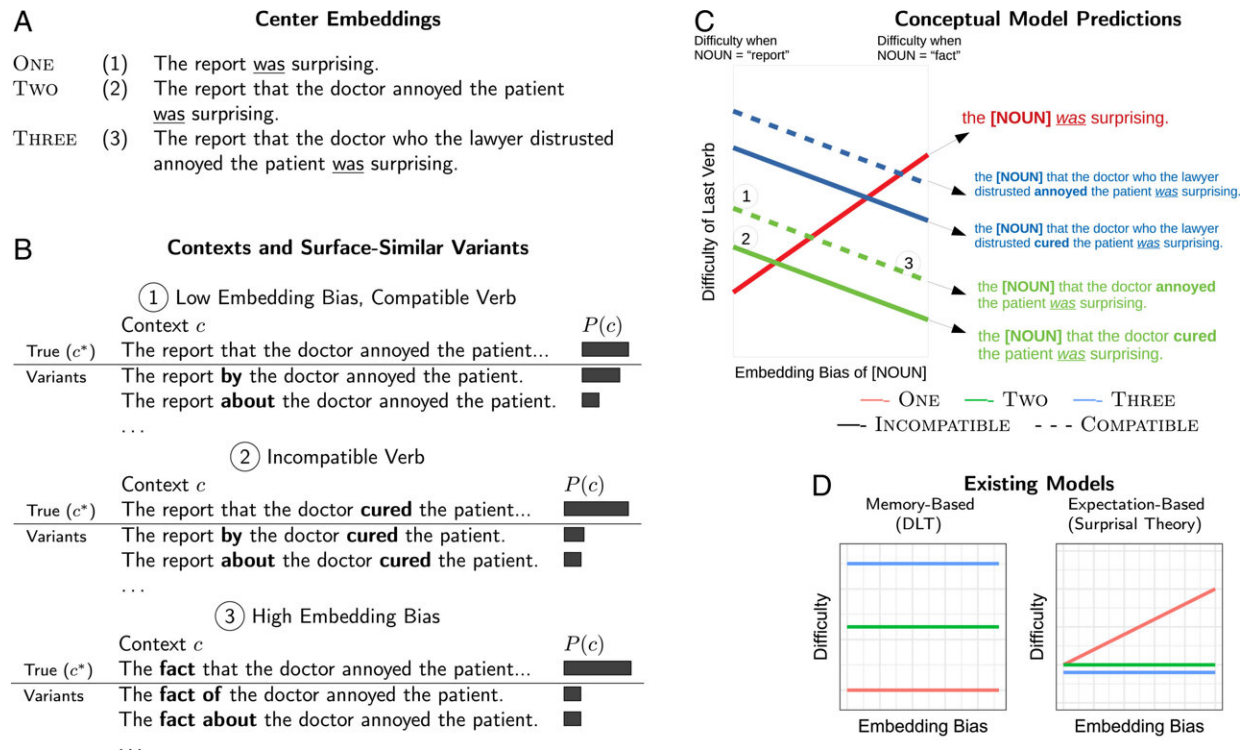
(A) Sentences exhibiting nested recursive structures ("center embeddings"), with one, two, or three noun–verb pairs. Difficulty differences between the sentences manifest as reading time differences on the underlined verbs. (B) Varying the probability of true contexts and surface-similar variants: We show a context requiring a verb in order to form a complete sentence, with surface-similar variants that do not require a verb in order to be complete (item 1). The relative prior probability P(c) of the true context and the surface-similar variants can be modulated by changing the identity of the second-to-last verb (item 2) or the first noun (item 3). Our model predicts that the final verb is easier to process when the true context has higher prior probability when compared to the variants. (C) Conceptual predictions of our model for difficulty at the last verb: First, higher difficulty is predicted for the Three condition (blue) than the Two condition (green). Second, higher difficulty is predicted when the second-to-last verb is

semantically compatible ("annoyed" as opposed to "cured") with the first noun. Third, higher difficulty is predicted when the first noun has a lower embedding bias ("report" as opposed to "fact"). In the One condition, an effect in the opposite direction is predicted. (D) Predictions are sharply different from existing theories: The DLT predicts that increasing levels of embedding should increase difficulty, but predicts no effects of embedding bias or compatibility. Surprisal theory predicts that embedding bias impacts difficulty in the One condition, but not in the other conditions. Credit: *Proceedings of the National Academy of Sciences* (2022). DOI: 10.1073/pnas.2122602119

In 2020, Futrell first devised a theory unifying these two models. He argued that limits in memory don't affect only retrieval in sentences with embedded clauses but plague all language comprehension; our memory limitations don't allow us to perfectly represent sentence contexts during language comprehension more generally.

Thus, according to this unified model, memory constraints can create a new source of difficulty in anticipation. We can have difficulty anticipating an upcoming word in a sentence even if the word should be easily predictable from context—in case that the sentence context itself is difficult to hold in memory. Consider, for example, a sentence beginning with the words "Bob threw the trash..." we can easily anticipate the final word—"out." But if the sentence context preceding the final word is more complex, difficulties in expectation arise: "Bob threw the old trash that had been sitting in the kitchen for several days [out]."

Researchers quantify comprehension difficulty by measuring the time it takes readers to respond to different comprehension tasks. The longer the response time, the more challenging the comprehension of a given sentence. Results from prior experiments showed that Futrell's unified account predicted readers' comprehension difficulties better than the two

older models. But his model didn't identify which parts of the sentence we tend to forget—and how exactly this failure in memory retrieval obfuscates comprehension.

Hahn's new study fills in these gaps. In the new paper, the [cognitive scientists](link) from MIT joined Futrell to propose an augmented model grounded in a new coherent theoretical framework. The new model identifies and corrects missing elements in Futrell's unified account and provides new fine-tuned predictions that better match results from empirical experiments.

As in Futrell's original model, the researchers begin with the idea that our mind, due to memory limitations, doesn't perfectly represent the sentences we encounter. But to this they add the theoretical principle of cognitive efficiency. They propose that the mind tends to deploy its limited memory resources in a way that optimizes its ability to accurately predict new word inputs in sentences.

This notion leads to several empirical predictions. According to one key prediction, readers compensate for their imperfect memory representations by relying on their knowledge of the statistical co-occurrences of words in order to implicitly reconstruct the sentences they read in their minds. Sentences that include rarer words and phrases are therefore harder to remember perfectly, making it harder to anticipate upcoming words. As a result, such sentences are generally more challenging to comprehend.

To evaluate whether this prediction matches our linguistic behavior, the researchers utilized GPT-2, an AI natural language tool based on neural network modeling. This machine learning tool, first made public in 2019, allowed the researchers to test the model on large-scale text data in a way that wasn't possible before. But GPT-2's powerful language modeling capacity also created a problem: In contrast to humans,

GPT-2's immaculate memory perfectly represents all the words in even very long and complex texts that it processes.

To more accurately characterize human language comprehension, the researchers added a component that simulates human-like limitations on memory resources—as in Futrell's original model—and used machine learning techniques to optimize how those resources are used—as in their new proposed model. The resulting model preserves GPT-2's ability to accurately predict words most of the time, but shows human-like breakdowns in cases of sentences with rare combinations of words and phrases.

"This is a wonderful illustration of how modern tools of machine learning can help develop cognitive theory and our understanding of how the mind works," says Gibson. "We couldn't have conducted this research here even a few years ago."

The researchers fed the machine learning model a set of sentences with complex embedded clauses such as, "The report that the doctor who the lawyer distrusted annoyed the patient was surprising." The researchers then took these sentences and replaced their opening nouns—"report" in the example above—with other nouns, each with their own probability to occur with a following clause or not.

Some nouns made the sentences to which they were slotted easier for the AI program to "comprehend." For instance, the model was able to more accurately predict how these sentences end when they began with the common phrasing "The fact that" than when they began with the rarer phrasing "The report that."

The researchers then set out to corroborate the AI-based results by conducting experiments with participants who read similar sentences. Their response times to the comprehension tasks were similar to that of

the model's predictions. "When the sentences begin with the words 'report that,'" people tended to remember the sentence in a distorted way," says Gibson. The rare phrasing further constrained their memory and, as a result, constrained their comprehension.

These results demonstrate that the new model out-rivals existing models in predicting how humans process language.

Another advantage the model demonstrates is its ability to offer varying predictions from language to language. "Prior models knew to explain why certain language structures, like sentences with embedded clauses, may be generally harder to work with within the constraints of memory, but our new model can explain why the same constraints behave differently in different languages," says Levy.

"Sentences with center-embedded clauses, for instance, seem to be easier for native German speakers than native English speakers, since German speakers are used to reading sentences where subordinate clauses push the verb to the end of the sentence."

According to Levy, further research on the model is needed to identify causes of inaccurate sentence representation other than embedded clauses. "There are other kinds of 'confusions' that we need to test." Simultaneously, Hahn adds, "the model may predict other 'confusions' which nobody has even thought about. We're now trying to find those and see whether they affect human comprehension as predicted."

Another question for future studies is whether the new model will lead to a rethinking of a long line of research focusing on the difficulties of sentence integration: "Many researchers have emphasized difficulties relating to the process in which we reconstruct language structures in our minds," says Levy. "The new model possibly shows that the difficulty relates not to the process of mental reconstruction of these sentences, but

to maintaining the mental representation once they are already constructed. A big question is whether or not these are two separate things."

One way or another, adds Gibson, "this kind of work marks the future of research on these questions."

More information: Michael Hahn et al, A resource-rational model of human processing of recursive linguistic structure, *Proceedings of the National Academy of Sciences* (2022). DOI: 10.1073/pnas.2122602119

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology