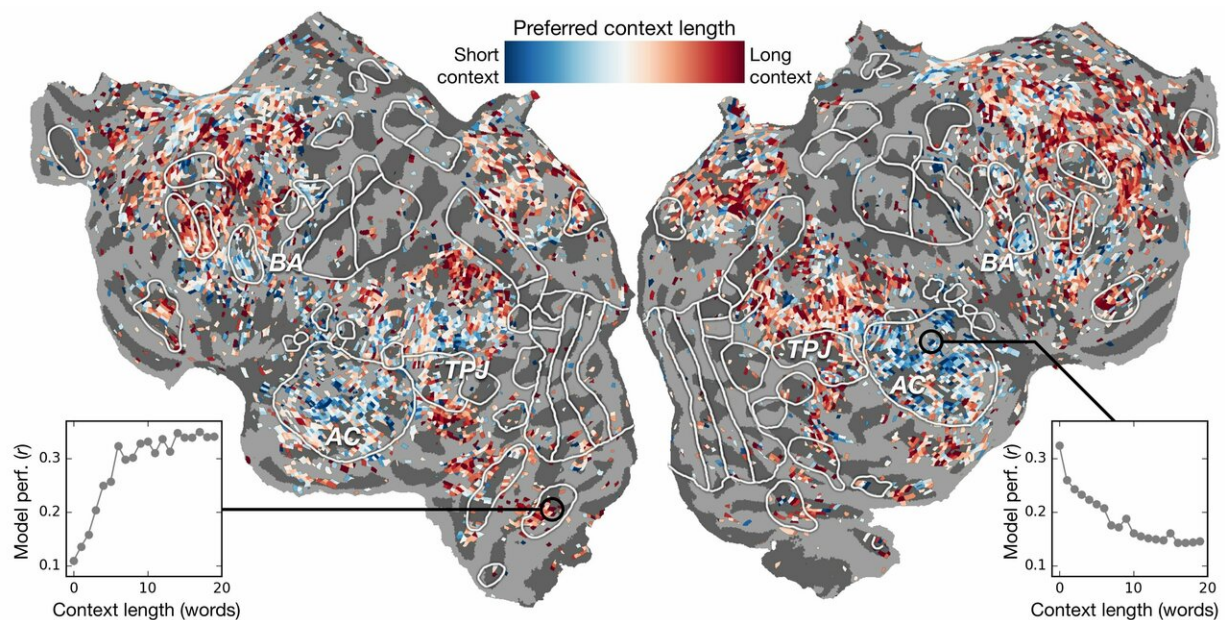


# Brain-inspired AI inspires insights about the brain (and vice versa)

March 20 2019, by Aaron Dubrow



Context length preference across cortex. An index of context length preference is computed for each voxel in one subject and projected onto that subject's cortical surface. Voxels shown in blue are best modeled using short context, while red voxels are best modeled with long context. Credit: Huth lab, UT Austin

Can artificial intelligence (AI) help us understand how the brain understands language? Can neuroscience help us understand why AI and neural networks are effective at predicting human perception?

Research from Alexander Huth and Shailee Jain from The University of Texas at Austin (UT Austin) suggests both are possible.

In a [paper](#) presented at the 2018 Conference on Neural Information Processing Systems (NeurIPS), the scholars described the results of experiments that used [artificial neural networks](#) to predict with greater accuracy than ever before how different areas in the brain respond to specific words.

"As words come into our heads, we form ideas of what someone is saying to us, and we want to understand how that comes to us inside the brain," said Huth, assistant professor of Neuroscience and Computer Science at UT Austin. "It seems like there should be systems to it, but practically, that's just not how [language](#) works. Like anything in biology, it's very hard to reduce down to a simple set of equations."

The work employed a type of recurrent neural network called long short-term memory (LSTM) that includes in its calculations the relationships of each word to what came before to better preserve context.

"If a word has multiple meanings, you infer the meaning of that word for that particular sentence depending on what was said earlier," said Jain, a Ph.D. student in Huth's lab at UT Austin. "Our hypothesis is that this would lead to better predictions of brain activity because the brain cares about context."

It sounds obvious, but for decades neuroscience experiments considered the response of the brain to individual words without a sense of their connection to chains of words or sentences. (Huth describes the importance of doing "real-world neuroscience" in a March 2019 [paper](#) in the *Journal of Cognitive Neuroscience*.)

In their work, the researchers ran experiments to test, and ultimately

predict, how different areas in the brain would respond when listening to stories (specifically, the Moth Radio Hour). They used data collected from fMRI (functional magnetic resonance imaging) machines that capture changes in the blood oxygenation level in the brain based on how active groups of neurons are. This serves as a correspondent for where language concepts are "represented" in the brain.

Using powerful supercomputers at the Texas Advanced Computing Center (TACC), they trained a language model using the LSTM method so it could effectively predict what word would come next—a task akin to Google auto-complete searches, which the [human mind](#) is particularly adept at.

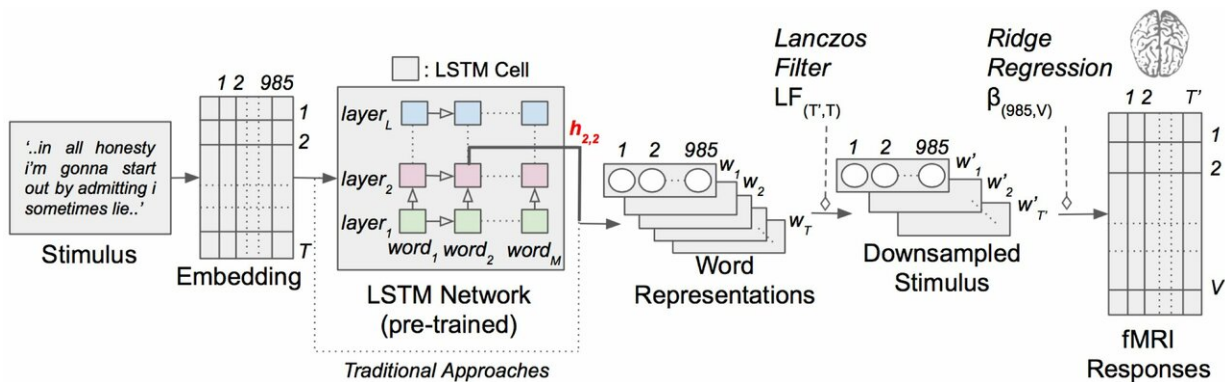
"In trying to predict the next word, this model has to implicitly learn all this other stuff about how language works," said Huth, "like which words tend to follow other words, without ever actually accessing the brain or any data about the brain."

Based on both the language model and fMRI data, they trained a system that could predict how the brain would respond when it hears each word in a new story for the first time.

Past efforts had shown that it is possible to localize language responses in the brain effectively. However, the new research showed that adding the contextual element—in this case up to 20 words that came before—improved brain activity predictions significantly. They found that their predictions improve even when the least amount of context was used. The more context provided, the better the accuracy of their predictions.

"Our analysis showed that if the LSTM incorporates more words, then it gets better at predicting the next word," said Jain, "which means that it must be including information from all the words in the past."

The research went further. It explored which parts of the brain were more sensitive to the amount of context included. They found, for instance, that concepts that seem to be localized to the auditory cortex were less dependent on context.



Contextual language encoding model with narrative stimuli. Each word in the story is first projected into a 985-dimensional embedding space. Sequences of word representations are then fed into an LSTM network that was pre-trained as a language model. Credit: Huth lab, UT Austin

"If you hear the word dog, this area doesn't care what the 10 words were before that, it's just going to respond to the sound of the word dog", Huth explained.

On the other hand, brain areas that deal with higher-level thinking were easier to pinpoint when more context was included. This supports theories of the mind and language comprehension.

"There was a really nice correspondence between the hierarchy of the artificial network and the hierarchy of the brain, which we found interesting," Huth said.

Natural language processing—or NLP—has taken great strides in recent years. But when it comes to answering questions, having natural conversations, or analyzing the sentiments in written texts, NLP still has a long way to go. The researchers believe their LSTM-developed language model can help in these areas.

The LSTM (and [neural networks](#) in general) works by assigning values in high-dimensional space to individual components (here, words) so that each component can be defined by its thousands of disparate relationships to many other things.

The researchers trained the language model by feeding it tens of millions of words drawn from Reddit posts. Their system then made predictions for how thousands of voxels (three-dimensional pixels) in the brains of six subjects would respond to a second set of stories that neither the model nor the individuals had heard before. Because they were interested in the effects of context length and the effect of individual layers in the neural network, they essentially tested 60 different factors (20 lengths of context retention and three different layer dimensions) for each subject.

All of this leads to computational problems of enormous scale, requiring massive amounts of computing power, memory, storage, and data retrieval. TACC's resources were well suited to the problem. The researchers used the Maverick supercomputer, which contains both GPUs and CPUs for the computing tasks, and Corral, a storage and data management resource, to preserve and distribute the data. By parallelizing the problem across many processors, they were able to run the computational experiment in weeks rather than years.

"To develop these models effectively, you need a lot of training data," Huth said. "That means you have to pass through your entire dataset every time you want to update the weights. And that's inherently very

slow if you don't use parallel resources like those at TACC."

If it sounds complex, well—it is.

This is leading Huth and Jain to consider a more streamlined version of the system, where instead of developing a language prediction model and then applying it to the brain, they develop a model that directly predicts brain response. They call this an end-to-end system and it's where Huth and Jain hope to go in their future research. Such a model would improve its performance directly on brain responses. A wrong prediction of brain activity would feedback into the model and spur improvements.

"If this works, then it's possible that this network could learn to read text or intake language similarly to how our brains do," Huth said. "Imagine Google Translate, but it understands what you're saying, instead of just learning a set of rules."

With such a system in place, Huth believes it is only a matter of time until a mind-reading system that can translate brain activity into language is feasible. In the meantime, they are gaining insights into both neuroscience and artificial intelligence from their experiments.

"The brain is a very effective computation machine and the aim of artificial intelligence is to build machines that are really good at all the tasks a brain can do," Jain said. "But, we don't understand a lot about the brain. So, we try to use [artificial intelligence](#) to first question how the [brain](#) works, and then, based on the insights we gain through this method of interrogation, and through theoretical neuroscience, we use those results to develop better artificial Intelligence.

"The idea is to understand cognitive systems, both biological and artificial, and to use them in tandem to understand and build better machines."



**More information:** Pawel J. Matusz et al, Are We Ready for Real-world Neuroscience?, *Journal of Cognitive Neuroscience* (2018). [DOI: 10.1162/jocn.e.01276](https://doi.org/10.1162/jocn.e.01276)

Provided by University of Texas at Austin

Citation: Brain-inspired AI inspires insights about the brain (and vice versa) (2019, March 20) retrieved 9 April 2024 from <https://techxplore.com/news/2019-03-brain-inspired-ai-insights-brain-vice.html>

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.