

Larger set of translations may shed light on an idiosyncratic Amazonian language

March 10 2016, by Peter Dizikes



The Pirahã reside along the Maici River, which branches off the Amazon in Brazil. MIT researchers are now making public the most extensive data set yet accumulated on the Pirahã language.

A heated controversy in linguistics in recent years involves a few

hundred people deep in the Amazonian rainforest: the Pirahã tribe of Northern Brazil. Their idiosyncratic language has raised questions about how widely human languages share certain characteristics.

Among the questions at issue is whether the Pirahã language contains recursion, a process through which [sentences](#) (and thus languages) can be expanded infinitely. Consider the sentence, "John wrote a book." We can add to it and, for instance, form noun phrases that contain multiple noun phrases themselves. Thus, "John and the man wrote a book" might be viewed as evidence of recursion.

Some linguists, including one who did some early fieldwork on the Pirahã, have argued that their language lacks recursion, making it anomalous among the world's tongues. Others, including some experts in MIT's Department of Linguistics, have disagreed with such claims. Many linguists view all languages as having universal affinities that help us understand what is unique about human language.

Now a newly published study co-authored by scholars in MIT's Department of Brain and Cognitive Sciences (BCS) has made public the most extensive data set yet accumulated on the Pirahã and reached an equivocal conclusion: The findings, they say, make it possible that the Pirahã language lacks recursion, without ruling out the possibility that recursion does exist in the tongue.

"We think it's consistent with there being no recursion, but we can't say for sure," says Edward Gibson, a professor in BCS and a co-author of the paper. "It's plausible."

To reach a more definitive conclusion, Gibson believes, "We would need so much more data." The current study is based on 1,100 sentences translated from Pirahã—more than ever assembled previously but not a large amount by the standards he has used when analyzing other

languages. The scholars are making the corpus available to anyone for future research.

"It's not just about injecting the data," Gibson says, suggesting the existence of the corpus means scholars will be less inclined to "talk about an example or two and then have very broad arguments ... We want a way of making the raw data available to everyone, so anyone can make their own conclusions based on open access."

Still, even given the larger dataset, the question of whether recursion is absent from the Pirahã language also hinges on the interpretations of scholars who have done fieldwork among the tribe—including Daniel L. Everett, a linguist and the dean of arts and sciences at Bentley University, who is a co-author of the paper and whose early field research has generated the sometimes intense debate on the subject.

Disappearing people

The paper, "A Corpus Investigation of Syntactic Embedding in Piraha," was published last week in the journal *PLOS One*. The co-authors are Gibson; Everett; Richard Futrell, a PhD student in BCS; Laura Stearns, a research assistant in BCS; and Steven Piantadosi, a professor of cognitive science at the University of Rochester.

In the 1970s, Everett lived for an extended time among the Pirahã, who reside along the Maici River, which branches off the Amazon. He published a doctoral dissertation on them in the 1980s, but the controversy over the Pirahã language did not explode until 2005, when Everett published a scholarly article making the case for the language's unique properties, including its apparent lack of recursion.

For the current paper, the researchers assembled the corpus by combining the transcriptions of 17 stories told in Pirahã. Those stories

were transcribed by Everett and Steve Sheldon, a missionary who lived among the Pirahã people in the 1970s.

Using the 1,100 sentences in the corpus, the researchers analyzed each one to see if they could find examples of several types of "recursive embedding." For instance, they looked for examples of "center embedding," in which a clause is inserted into the middle of a sentence. No examples of it were present in the Pirahã dataset.

However, the issue of whether some Pirahã sentences in the corpus display recursion may also be a matter of translation and interpretation. For example, the use of conjunction in a sentence can lead to boundless new forms of sentences, by joining noun phrases such as "Joe and Sue went to the market."

The set of 1,100 sentences contains five cases that could suggest use of conjunction. In three of those cases, the co-authors believe, this simply represents a close juxtaposition of noun phrases without any special linking of them.

In the other two cases, Everett's translations of the sentences differ from those of Sheldon. Everett believes there is no conjunction in those sentences, although Sheldon's original translations suggested there was.

For instance, a Pirahã sentence transliterated as "ti xaigia ao ogi gio ai hi ahapita" was interpreted by Sheldon to mean, "Well, then I and the big Brazilian woman disappeared." The conjunction of "I and the big Brazilian woman" could be an example of recursion. But Everett believes a better translation is: "Well, [with respect to me], the very big foreigner went away again." And that sentence has no recursion.

"Let's get everything out there"

The co-authors say they do not expect the current paper to end all debate but rather to help scholars see the comparative frequency (or lack thereof) by which sentences that are even suggestive of recursion appear in Pirahã. In addition, the translations provide more examples of the language in everyday use.

As Futrell notes, "I think the approach of analyzing the data in a rigorous way is more important than whatever conclusion we've come to about Pirahã recursion." Futrell also suggests that "the right way to think about universalism is not in terms of all languages [containing] X or Y, but rather that there's some probability distribution over all the properties languages can have."

Other scholars say it is good to have more data available. Tom Wasow, a professor emeritus of linguistics at Stanford University, who has seen the paper, calls it a "careful, systematic" examination of the corpus, adding: "Probably the most important contribution of the paper is making the corpus publicly available, so that investigators can check for themselves whether they see anything that looks like recursion."

Still, Wasow, who taught Futrell when Futrell was an undergraduate, suggests that even resolving the matter of whether or not Pirahã contains recursion would not negate certain claims about the global affinities of [human languages](#).

"Suppose one, or even a few dozen, of the thousands of languages in the world lack recursion," Wasow states. "Should linguists therefore conclude that recursion is not a general property of language? I don't think so." By analogy, Wasow adds, "When biologists discovered the platypus, they did not abandon the generalization that mammals give live birth to their offspring; rather they recognized that it is a characteristic of almost all species of mammals."

David Pesetsky, the Ferrari P. Ward Professor of Modern Languages and Linguistics at MIT, and head of the MIT Department of Linguistics and Philosophy, has previously published articles criticizing Everett's findings, and believes an alternate analytical approach will be needed to reach any conclusions about the structure of Pirahã.

Regarding recursion specifically, for instance, Pesetsky notes that a graduate student at the University of British Columbia, Raiane Salles, has conducted recent fieldwork among the Pirahã unearthing potential evidence of "possessor recursion"—which creates phrases that could expand infinitely, such as "the foreigner's parent's dog" and "Migixoi's husband's mother's clothes."

While the current paper has not changed Pesetsky's position, he notes that he does appreciate having a corpus of Pirahã made public.

Gibson says he is willing to see the scholarly debate unfold, based on access to the larger data set: "Let's get everything out there, and we can all talk about it."

More information: Richard Futrell et al. A Corpus Investigation of Syntactic Embedding in Pirahã, *PLOS ONE* (2016). [DOI: 10.1371/journal.pone.0145289](https://doi.org/10.1371/journal.pone.0145289)

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

Citation: Larger set of translations may shed light on an idiosyncratic Amazonian language (2016, March 10) retrieved 25 April 2024 from <https://phys.org/news/2016-03-larger->

idiosyncratic-amazonian-language.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.