Language structure arises from balance of clear and effective communication

October 15 2012

When learning a new language, we automatically organize words into sentences that will be both clearly understood and efficient (quick) to communicate. That's the finding of a new study reported today in the *Proceedings of the National Academy of Sciences* (*PNAS*) which challenges opposing theories on why and how languages come to be organized the way they are.

With more than 5000 languages in the world, it would be easy to assume all vary endlessly, but, in fact, there is great commonality: languages follow only a few recurrent patterns. These commonalities are called "language universals," a notion suggested in the 1960's by Noam Chomsky and Joseph Greenberg. A team of researchers from the University of Rochester and Georgetown University Medical Center set out to investigate how these language universals come to be.

Linguists and cognitive scientists have opposing ideas on how a language is developed and shaped. Some believe that languages all derived from a common ancestor; others think that languages vary quite widely and universals do not exist at all. Some have suggested that language universals are an arbitrary evolutionary outcome. The position of the Rochester-Georgetown team is that the human mind shapes a language, even while learning it, based on the need for robust and effective information transfer.

"The thousands of natural languages in our world only have a couple of formats in which they appear, and we are good at understanding and

learning languages that have just these formats. Otherwise we could never succeed in learning something so complicated as human languages ," says one of the study's authors, Elissa L. Newport, Ph.D., a professor in the department of neurology at Georgetown University Medical Center.

A member of the National Academy of Sciences, Newport is also director of the Center for Brain Plasticity and Recovery, a joint program of Georgetown University and MedStar National Rehabilitation Network.

The study was conducted by Rochester graduate researcher Maryia Fedzechkina in a collaboration with Newport and T. Florian Jaeger, Ph.D., Wilmot Assistant Professor of the Sciences in the department of brain and cognitive sciences at Rochester.

According to Jaeger, the question that motivated their study is "whether subtle biases towards language codes that facilitate efficient information transfer operate during language acquisition, causing learners to deviate from the input they receive, slowly changing languages over generations."

For the study, participants were taught a miniature artificial language made of nonsensical words including 8 verbs and 15 nouns (e.g. kliedum, slergin, zub and zamper). The volunteers were shown videos while hearing sentences from the language, and spent several days learning the language. But the language they were exposed to was organized unlike any natural language. While many languages have prefixes or suffixes on nouns to indicate subject or object – a property called case-marking – their artificial languages contained case marking on only about 60 percent of the nouns, and this sprinkling of case markers did not follow any of the principles that appear in real languages regarding when a case marker would be most likely or most helpful.

The researchers wanted to know if the participants would "fix" the language on their own.

"English, as well as a lot of other languages, use subject-verb-object word order to indicate roles, such as 'the boy kicked the wall'," Fedzechkina explains. "Other languages use case markings on nouns. Of particular interest is that some languages such as Korean and Japanese use case markings precisely when the sentence would be easily misunderstood without them."

The participants in this study spoke only English—they were not familiar with other languages that use case markings. Yet after days of training in the made-up language, when participants were asked to speak new sentences to describe a video in which one person is doing something to another, they deviated from what was taught. They added case markings to nouns precisely when it made the sentence clearer in determining the object and subject.

"They fixed the sentence structure," Newport says.

"The study's results support the idea that we apply a 'language universal' that services clear yet efficient communication," Newport says. She adds, what makes this study stronger is what the volunteers didn't do.

"They could have removed all the inconsistent words and produced a language with none of the case markings," she explains. "They could have reproduced what we gave them, keeping the meaning of the markers and sentences uncertain. Or they could have added a case marker in every sentence, to mark every object, which would have made the sentences clear but also long and inefficient."

But they didn't.

"What they did was add the case marker only in cases where the meaning would have otherwise been ambiguous and confusing." Newport concludes, "We found that when you make up languages that violate universals, people will change the language, moving it toward universal principles."

Provided by Georgetown University Medical Center