Researchers find common cognitive foundation for child language development and language evolution

July 27 2023

Credit: CC0 Public Domain

Cognitive and computer scientists at the University of Toronto, Universitat Pompeu Fabra and the Catalan Institution for Research and Advanced Studies have found child language development and the historical evolution of the world's languages share a common cognitive

foundation—a core knowledge base where patterns of children's language innovation can predict patterns of language evolution, and vice versa.

Published today in *Science*, the paper is a first-of-its-kind step toward a unified theory of the lexicon and the mind examined across timescales. The result may also help predict how a word's meaning may change in the future—across different languages, in language learners and in machine learning.

For the study, the team focused on a common form of human lexical creativity, or word coinage, known as word meaning extension—where people use known words to express something new instead of creating new words. For example, the word "mouse" in the historical evolution of English extends from its rodent meaning to refer to a portable computer device.

In language development, children as young as two years old can use the word "ball" to refer to "balloon," presumably because they haven't yet acquired the right word to describe "balloon," so they overextend the known word "ball" to express that new object.

"We investigated processes of word meaning extension across populations and within individuals, and at two very different timescales—in language change and evolution, which take over hundreds and thousands of years, and in child language development during the first few months and years of life," says last author Yang Xu, Associate Professor, Department of Computer Science, Cognitive Science Program, University of Toronto.

"We found that these diverse processes are fundamentally the same, and that the creation of new word meanings relies on a shared foundation of knowledge grounded in human experience."

First author Thomas Brochhagen, Assistant Professor, Department of Translation and Language Sciences, Universitat Pompeu Fabra says, "This possible relationship between individual learning and the evolution of languages in terms of how meaning is organized had not been demonstrated thus far, and our study does so on a large scale and in a generalized way."

For the study, the researchers built a computational model that takes pairs of concepts as input, such as "ball" versus "balloon" and "door" versus "key" and makes a prediction about how likely these concepts can be co-named under the same word.

To identify similarities between concepts, the model draws on four primary knowledge types grounded in human experience: visual, associative, taxonomic (how terms are organized in a hierarchy, like referring to an apple as a fruit), and affective (how pleasant and intense a term is, like "sunny").

The pair of concepts like "ball" and "balloon" would score high due to their similar visual features, whereas "door" and "key" would score high because they are thematically related or often occur together in daily scenarios. "Water" and "pencil" would have little similarity measured in any of the four knowledge types, so that pair would receive a low score. As a result, the model would predict they are unlikely to extend to each other.

The team found that the four knowledge types contributed to word meaning extension which indicates that word meaning extension relies on multifaceted and grounded knowledge based on people's perceptual, affective, and common-sense knowledge.

The researchers then performed a cross-predictive analysis using a model built exclusively from children's word meaning extension data to

predict successfully word meaning extension patterns from both language evolution and language change, and in the reverse.

The researchers also checked the robustness of these predictive models in languages other than English, verifying that the creation of new word meanings follows similar patterns in 1,400 different languages, including Spanish, Catalan, Basque, Galician, German, French, Portuguese, Dutch, Danish, Norwegian, Swahili, Arabic, Mandarin Chinese, Hindi and Korean.

Existing research on child overextension is typically discussed in the context of developmental psychology whereas word meaning extension in history is typically discussed in historical and computational linguistics.

"By building this connection between the two fields, we find a core knowledge engine that supports lexical creativity in word meaning extension, which is fundamentally important to human cognition and linguistic communication of emerging meanings," says Xu.

These computational models may also help facilitate and understand second language acquisition by interpreting errors that learners made in English and other languages, which could resemble how children and adults extend word meaning in their mother tongue.

Future research will further explore the origins and cognitive mechanisms of human lexical creativity, and the possibility of predicting new or emerging meaning in both human language development and machine learning systems.

"Developing a unified theory of the mind across timescales is a challenging undertaking—we don't have access to human minds dating back hundreds or thousands of years," says Xu. "Our study offers an

alternative way for exploring this unification through the lexicon—a creative product of the human mind and the system of word-meaning mappings, for which we have data available to us."

More information: Thomas Brochhagen, From language development to language evolution: A unified view of human lexical creativity, *Science* (2023). DOI: 10.1126/science.ade7981. www.science.org/doi/10.1126/science.ade7981

Provided by University of Toronto