Researchers use a computer model to explain how children integrate information during word learning

July 1 2021

Credit: Pixabay/CC0 Public Domain

Children learn a huge number of words in the early preschool years. A two-year-old might be able to say just a handful of words, while a five-

year-old is quite likely to know many thousands. How do children achieve this marvelous feat? The question has occupied psychologists for over a century: In countless carefully designed experiments, researchers titrate the information children use to learn new words. How children integrate different types of information has remained unclear.

"We know that children use a lot of different information sources in their social environment, including their own knowledge, to learn new words. But the picture that emerges from the existing research is that children have a bag of tricks that they can use", says Manuel Bohn, a researcher at the Max Planck Institute for Evolutionary Anthropology.

For example, if you show a child an object they already know—say a cup—as well as an object they have never seen before, the child will usually think that a word they never heard before belongs with the new object. Why? Children use information in the form of their existing knowledge of words (the thing you drink out of is called a "cup") to infer that the object that doesn't have a name goes with the name that doesn't have an object. Other information comes from the social context: children remember past interactions with a speaker to find out what they are likely to talk about next.

"But in the real world, children learn words in complex social settings in which more than just one type of information is available. They have to use their knowledge of words while interacting with a speaker. Word learning always requires integrating multiple, different information sources", Bohn continues. An open question is how children combine different, sometimes even conflicting, sources of information.

Predictions by a computer program

In a new study, a team of researchers from the Max Planck Institute for Evolutionary Anthropology, MIT, and Stanford University takes on this

issue. In a first step, they conducted a series of experiments to measure children's sensitivity to different information sources. Next, they formulated a computational cognitive model which details the way that this information is integrated.

"You can think of this model as a little computer program. We input children's sensitivity to different information, which we measure in separate experiments, and then the program simulates what should happen if those information sources are combined in a rational way. The model spits out predictions for what should happen in hypothetical new situations in which these information sources are all available", explains Michael Henry Tessler, one of the lead-authors of the study.

In a final step, the researchers turned these hypothetical situations into real experiments. They collected data with two- to five-year-old children to test how well the predictions from the model line up with real-world data. Bohn sums up the results: "It is remarkable how well the rational model of information integration predicted children's actual behavior in these new situations. It tells us we are on the right track in understanding from a mathematical perspective how children learn language."

Language learning as a social inference problem

How does the model work? The algorithm that processes the different information sources and integrates them is inspired by decades of research in philosophy, developmental psychology, and linguistics. At its heart, the model looks at language learning as a social inference problem, in which the child tries to find out what the speaker means—what their intention is. The different information sources are all systematically related to this underlying intention, which provides a natural way of integrating them.

Additionally, the model also specifies what changes as children get older.

Over development, children become more sensitive to the individual information sources, and yet the social reasoning process that integrates the information sources remains the same.

"The virtue of computational modeling is that you can articulate a range of alternative hypotheses—alternative models—with different internal wiring to test if other theories would make equally good or better predictions. In some of these alternatives, we assumed that children ignore some of the information sources. In others, we assumed that the way in which children integrate the different information sources changes with age. None of these alternative models provided a better explanation of children's behavior than the rational integration model", explains Tessler.

The study offers several exciting and thought-provoking results that inform our understanding of how children learn language. Beyond that, it opens up a new, interdisciplinary way of doing research. "Our goal was to put formal models in a direct dialogue with experimental data. These two approaches have been largely separated in child development research", says Manuel Bohn. The next steps in this research program will be to test the robustness of this theoretical model. To do so, the team is currently working on experiments that involve a new set of information sources to be integrated.

The study is published in *Nature Human Behaviour*.

More information: How young children integrate information sources to infer the meaning of words, *Nature Human Behaviour* , DOI: 10.1038/s41562-021-01145-1 , www.nature.com/articles/s41562-021-01145-1