

Order matters: Developmentally plausible acquisition of lexical categories

Philip Huebner

University of Illinois, Urbana-Champaign, Urbana, Illinois, United States

Jon Willits

University of Illinois at Urbana-Champaign, Champaign, Illinois, United States

Abstract

One proposal for how children acquire syntactic and semantic lexical categories is by inducing them from their distributional signatures in speech. Because the language children are exposed to gradually increases in complexity as they get older, it is possible that inducing lexical categories from initially simplified speech supports acquisition. We set out to test this hypothesis using a simple recurrent neural network trained to predict 5 million words of child-directed speech from the American-English portion of the CHILDES database. Evaluation of learned representations showed that models trained in order in which children actually experience language performed better on a semantic, but not syntactic, categorization task. To understand why, we examined how the models encoded words during the earliest stages of training. Our results are relevant to important questions in language acquisition, such as the role of early experiences in organizing children's linguistic representations.