

A computational approach to understanding how infants perceive language

29 January 2021



Credit: CC0 Public Domain

Languages differ in the sounds they use. The Japanese language, for example, does not distinguish between 'r' and 'l' sounds as in 'rock' versus 'lock.' Remarkably, infants become attuned to the sounds of their native language before they learn to speak. One-year-old babies, for example, less readily distinguish between 'rock' and 'lock' when living in an environment where Japanese, rather than English, is spoken.

Influential scientific accounts of this early phonetic learning phenomenon initially proposed that infants group sounds into native vowel- and consonant-like phonetic categories through a statistical clustering mechanism known as distributional learning.

The idea that infants learn consonant- and vowel-like phonetic categories has been challenged, however, by a new study published this week in the *Proceedings of the National Academy of Sciences*.

In the study, a multi-institutional team of cognitive scientists and computational linguists have

introduced a quantitative modeling framework that is based on a large-scale simulation of the [language](#) learning process in infants. Using computationally efficient machine learning techniques, this approach allows learning mechanisms to be systematically linked to testable predictions regarding infants' attunement to their native language.

"Hypotheses about *what* is being learned by infants have traditionally driven researchers' attempts to understand this surprising phenomenon," says Thomas Schatz, a postdoctoral associate in the University of Maryland of Maryland Institute for Advanced Computer Studies (UMIACS) who was lead author of the study. "We propose to start from hypotheses about *how* infants might learn."

In addition to Schatz, the study's authors include Naomi Feldman, an associate professor of linguistics at the University of Maryland with an appointment in UMIACS; Sharon Goldwater, a professor in the Institute for Language, Cognition and Computation at the University of Edinburgh's School of Informatics; Xuân-Nga Cao, a research engineer at Ecole Normale Supérieure (ENS) in Paris and co-founder of the Langinnov and Gazouyi startups; and Emmanuel Dupoux, a professor who directs the Cognitive Machine Learning team at ENS.

For their study, the researchers simulated the learning process in infants by training a computationally efficient clustering algorithm on realistic speech input. The algorithm was fed spectrogram-like auditory features sampled at regular time intervals that were obtained from naturalistic speech recordings in a target language. In this study, American English and Japanese were the two languages used.

This yielded a candidate model for the early phonetic knowledge of, say, a Japanese infant, the researchers say. Next, they asked two questions of

the trained models. Could they explain the observed differences in how Japanese- and English-learning infants discriminate speech sounds? And, did the models learn vowel- and consonant-like phonetic categories?

The dominant scientific accounts of early phonetic learning would have expected the answers to these questions to match (either both should be 'yes' or both should be 'no'). The researchers found that the answer to the first question was positive: Their models did account for infants' observed behavior, in particular for the Japanese infants' difficulty with distinguishing words like 'rock' and 'lock.' The answer to the second question, however, was negative: The models were found to have learned speech units too brief and acoustically variable to correspond to vowel- and consonant-like phonetic categories.

These results suggest a striking reinterpretation of the existing literature on early phonetic learning. Difficulties in scaling up distributional learning of phonetic categories to realistic learning conditions may be better interpreted as questioning the idea that *what* infants learn are phonetic categories, rather than the idea that *how* infants learn is through pure distributional learning (the traditional interpretation).

Cognitive science has not traditionally made use of such large-scale modeling, says Schatz, but recent advances in computing power, large datasets, and machine-learning algorithms make this approach more feasible than ever before.

Schatz and Feldman are part of the Computational Linguistics and Information Processing (CLIP) Laboratory in UMIACS, where Feldman is the current director. The robust computing resources in the CLIP lab and the Cognitive Machine Learning lab in Paris were instrumental to the research project, Feldman says.

In conclusion, the researchers believe their computationally-based modeling approach—together with ongoing efforts in the field to collect empirical data on a large scale, such as large-scale recordings of infants' learning environments at home and large-scale assessment

of [infants](#)' learning outcomes—opens the path toward a much deeper understanding of early language acquisition.

More information: Thomas Schatz et al, Early phonetic learning without phonetic categories: Insights from large-scale simulations on realistic input, *Proceedings of the National Academy of Sciences* (2021). [DOI: 10.1073/pnas.2001844118](https://doi.org/10.1073/pnas.2001844118)

Provided by University of Maryland

APA citation: A computational approach to understanding how infants perceive language (2021, January 29) retrieved 12 July 2022 from <https://medicalxpress.com/news/2021-01-approach-infants-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.