Learning word-referent mappings and concepts from raw inputs

Wai Keen Vong
NYU, New York, New York, United States

Brenden Lake
NYU, New York, New York, United States

Abstract

How do children learn correspondences between the language and the world from noisy, ambiguous, naturalistic input? One hypothesis is via cross-situational learning: tracking words and their possible referents across multiple situations allows learners to disambiguate correct word-referent mappings (Yu and Smith, 2007). While previous models of cross-situational word learning operate on highly simplified representations, recent advances in multimodal learning have shown promise as richer models of cross-situational word learning to enable learning the meanings of words from raw inputs. Here, we present a neural network model of cross-situational word learning that leverages some of these ideas and examine its ability to account for a variety of empirical phenomena from the word learning literature.