

Estimating Lexical-Semantic Networks from Verbal Fluency Data in 5-8 Year Olds

AJ Jain

UT Southwestern Medical Center, Dallas, Texas, United States

Martin Zettersten

University of California San Diego, La Jolla, California, United States

Casey Lew-Williams

Princeton University, Princeton, New Jersey, United States

Abstract

As children learn language, they organize their knowledge in lexical-semantic networks. Comparing pre-existing methods of accessing underlying networks, we examine developmental verbal fluency in 5-8-year-olds ($N=37$, mean age=80.27 mos) across two different prompt types – taxonomic (animals and foods) vs. (location-based) thematic prompts (zoo and grocery store) – using two different graph-theoretic estimation strategies: random-walk modeling (e.g., U-INVITE) and GloVe word embeddings. We observed several consistencies: taxonomic prompts elicited more words than thematic prompts (linear mixed-effects model: $t(35) = 3.16$, $p<0.05$); networks expanded with age (U-INVITE, $t(35) = 4.26$, $p<0.05$; ESN, $t(35) = 4.53$, $p<0.05$); and structures spread out (versus clustering densely). However, key differences emerged. Random-walk networks uncovered different highest-degree (most densely clustered) words depending on the prompt type (e.g., “dog” for animals, “monkey” for zoo). By contrast, networks based on word embeddings identified networks with very similar highest-degree words for animals and zoo. Hence, alternative assumptions informing method choices may result in distinct network estimates, with consequences for how we map the growth of lexical knowledge.