

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How to Build a Toddler Lexical Network

Permalink

<https://escholarship.org/uc/item/5hp3g2ss>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Weber, Jennifer
Colunga, Eliana

Publication Date

2022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

How to Build a Toddler Lexical Network

Jennifer Weber

University of Colorado, Boulder, Boulder, Colorado, United States

Eliana Colunga

University of Colorado Boulder, Boulder, Colorado, United States

Abstract

Understanding child language development requires accurately representing children's lexicons. However, past work modeling children's lexical-semantic structure typically utilized adult norms and corpora. The present work uses Word2Vec embeddings trained on a newly-created toddler-directed language corpus. Distributional approaches like Word2Vec calculate similarities taking into account not just when words occur together, but also when words occur in similar contexts. A network created from Word2Vec embeddings showed higher accuracy in predicting normed word acquisition from 16 to 30 months using network centrality measures, when compared to a network created using sliding window co-occurrences. We also compared predictions from the Word2Vec toddler network, a network created by training Word2Vec on typical adult input, and a model trained using both corpora. The toddler-only network outperformed the other two, indicating the importance of selecting language sources that reflect the population of interest. The present results reveal a promising new direction in understanding toddler word learning.