UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Word Prediction in Context: An Empirical Investigation of Core Vocabulary

Permalink

https://escholarship.org/uc/item/0953v3g2

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors

Wang, Andrew
De Deyne, Simon
McKague, Meredith
et al.

Publication Date

2023

Copyright Information

Peer reviewed

Word Prediction in Context: An Empirical Investigation of Core Vocabulary

Andrew Wang
The University of Melbourne, Melbourne, VIC, Australia

Simon De Deyne
University of Melbourne, Melbourne, VIC, Australia

Meredith McKague
University of Melbourne, Melbourne, VIC, Australia

Andrew Perfors
University of Melbourne, Melbourne, VIC, Australia

Abstract

Core vocabulary is a topic of huge interest in linguistics and has been studied from a wide variety of perspectives, such as language learning, dictionary studies, and cross-linguistically. In many of these conceptions, word frequency is widely considered the conventional measure of a word's coreness; however, this approach overlooks important aspects of mental representation like centrality in an associative semantic network. In this experiment, we compare different approaches to defining core words in a task that involves predicting missing words in sentences. Results showed that core words (regardless of definition) were easier to guess than non-core words, but that frequency-defined ones did not perform as well as expected given their higher predictability and the nature of the task. Analysis of incorrect responses also showed that people preferred to guess core words, simple synonyms, and words that are taxonomically related to the target. The findings suggest that how core vocabulary is defined depends in part on the nature of the task and that aspects of both mental representation and the linguistic environment play an important role.