UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Contextual Diversity and the Lexical Organization of Multiword Expressions

Permalink

https://escholarship.org/uc/item/1rx1552q

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

Authors

Senaldi, Marco Silvio Giuseppe
Titone, Debra
Johns, Brendan

Publication Date

2021

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

eScholarship.org Powered by the California Digital Library
University of California

Contextual Diversity and the Lexical Organization of Multiword Expressions

Marco Silvio Giuseppe Senaldi
McGill University, Montreal, Quebec, Canada

Debra Titone
McGill University, Montreal, Quebec, Canada

Brendan Johns
McGill University, Kingston, Ontario, Canada

Abstract

Corpus-based models of lexical strength have questioned the role of word frequency in lexical organization. Specifically, closer fits to lexical behavior data on single words have been obtained by measures of contextual diversity, which modifies frequency by ignoring word repetition in context, semantic diversity, which considers the semantic consistency of contextual word distribution, and socially-based semantic diversity, which encodes the communication patterns of individuals across discourses (Adelman, Brown & Quesada, 2006; Jones, Johns, & Recchia, 2012; Johns, in press). The present work aimed at determining if diversity drives lexical organization also at the level multiword units. Normative ratings of familiarity for 210 English idioms (Libben & Titone, 2008) were predicted from contextual, semantic and socially-based diversity measures computed from a 55-billion word corpus of Reddit comments. Results confirmed the superiority of diversity measures over word frequency, suggesting that multiword idiomatic phrases show similar lexical organization dynamics as single words.