

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

WAX: A Dataset for Word Association eXplanations

Permalink

<https://escholarship.org/uc/item/5rh0058r>

Authors

Liu, Chunhua

Cohn, Trevor

De Deyne, Simon

et al.

Publication Date

2023

Peer reviewed

# WAX: A Dataset for Word Association eXplanations

**Chunhua Liu**

The University of Melbourne, Melbourne, VIC, Australia

**Trevor Cohn**

The University of Melbourne, Melbourne, Australia

**Simon De Deyne**

University of Melbourne, Melbourne, VIC, Australia

**Lea Frermann**

University of Melbourne, Melbourne, Australia

## Abstract

Word associations are among the most common paradigms for studying the human mental lexicon. While their structure and underlying relation types have been widely studied, surprisingly little attention has been given to the question of why participants produce the observed associations. Answering this question would not only advance understanding of human cognition, but could also aid machines in learning and representing basic commonsense knowledge. Here, we introduce WAX: a large, crowd-sourced data set of 19K English word associations paired with human-generated free-text explanations and relation type labels. We present an efficient framework for eliciting associations together with explanations, and a comprehensive analysis of the emerging types of explanations and relations. We test language models in their ability to predict associations and generate explanations, demonstrating that models struggle to capture the diversity of human-produced associations, and suggesting WAX as a fertile resource for future research.