# UC Merced

**Proceedings of the Annual Meeting of the Cognitive Science Society**

**Title**

Evaluating word association-derived word embeddings on semantic analogies

**Permalink**

https://escholarship.org/uc/item/10g2h6qv

**Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

**Authors**

Lión, Rocío Belén
Quintero, Hernan
Valle-Lisboa, Juan C
et al.

**Publication Date**

2024

**Copyright Information**

Peer reviewed

# Evaluating word association-derived word embeddings on semantic analogies

**Rocío Lión**
Universidad de la República, Montevideo, Uruguay

**Hernan Quintero**
Universidad de la República, Montevideo, Uruguay

**Juan Valle-Lisboa**
Universidad de la República, Montevideo, Montevideo, Uruguay

**Alvaro Cabana**
Universidad de la República, Montevideo, Uruguay

## Abstract

Word embeddings trained on large scale text corpora are central to modern natural language processing and are also important as cognitive models and tools in psycholinguistic research (Pennington et al., 2014). An important alternative to these text-based models are embeddings derived from word association norms (De Deyne et al., 2019). Recently, these association-based embeddings have been shown to outperform text-based word embeddings of comparable complexity (such as GloVE, word2Vec & fastText) in semantic similarity rating tasks (Cabana et al., 2023; Richie & Bhatia, 2021). Here we evaluate English and Rioplatense Spanish association-based embeddings derived from the Small World of Words (SWOW) project on the Google Analogy set and the Bigger Analogy Test Set (Gladkova et al., 2016). We also developed a small analogy set that focuses on semantic relationships, such as event knowledge and category-exemplar relationships such as prototypicality. SWOW-derived word embeddings perform similarly as traditional text-based word embeddings in semantic analogies, and outform them in some categories. These results illustrate relevant similarities and differences between text-based and word association-derived embeddings.

References

Cabana, Á., Zugarramurdi, C., Valle-Lisboa, J. C., & De Deyne, S. (2023). The "Small World of Words" free association norms for Rioplatense Spanish. Behavior Research Methods. https://doi.org/10.3758/s13428-023-02070-z De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. Behavior Research Methods, 51(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7 Gladkova, A., Drozd, A., & Matsuoka, S. (2016). Analogy-based detection of morphological and semantic relations with word embeddings: What works and what doesn't. Proceedings of the NAACL Student Research Workshop, 8–15. https://doi.org/10.18653/v1/N16-2002 Richie, R., & Bhatia, S. (2021). Similarity Judgment Within and Across Categories: A Comprehensive Model Comparison—Richie—2021—Cognitive Science—Wiley Online Library. Cognitive Science, e13030. https://doi.org/10.1111/cogs.13030