ANITA B.ORG

# Modeling Free Associations Using Distributional Semantic Models and Spreading Activation

**Abhilasha A. Kumar**
Washington University in St Louis,
St Louis, MO 63130
+1-314-935-6524
abhilasha.kumar@wustl.edu

**David A. Balota**
Washington University in St Louis,
St Louis, MO 63130
+1-314-935-6524
dbalota@wustl.edu

**Simon De Deyne**
University of Melbourne,
Parkville VIC 3010, Australia
+61 3 9035 5511
simon.dedeyne@unimelb.edu.au

## 1. INTRODUCTION AND MOTIVATION

The past decade has seen an explosion of machine-learning based language models that can adequately perform complex language tasks like question answering, verbal analogies, etc. [1]. However, the temporal dynamics underlying how information becomes available over time in these models remains unknown, and although language models perform well in familiarity-based tasks like lexical decision, their performance has not been adequately assessed in production-based attentional tasks like free association (where participants are asked to produce the first word that comes to mind when given a cue). Free association represents a unique instance of retrieval from memory when the search process is relatively unconstrained and the task represents a universal phenomenon. Although some work has examined responses in this task, the underlying semantic representations that influence response *latencies* in this task have not been studied at all.

In the current research, we present a computational model that combines vector representations derived from three embedding models with a spreading activation framework [2] to analyze response proportions and latencies (RTs) in a continued free association task. Our work is based on a subset of the Small World of Words dataset [3], which contains primary, secondary, and tertiary responses to over 12,000 cue words, collected from over 88,000 participants including 62% who identified as female, 38% male, and less than 1% participants from other genders. Participants were from a wide variety of English-speaking countries, including USA, UK, Canada, and Australia, indicating the universality and inclusivity of free associations and their potential to map out search processes in populations worldwide.

We compare the predictive power of word2vec, fastText, and GloVe embeddings when combined with temporal parameters in explaining variance in producing the first response in the continued free association task, and also discuss how this research could be extended to examining propagation of ideas in specific social and cultural contexts.

## 2. RELATED WORK

### 2.1 WORD EMBEDDING MODELS

Word2vec, fastText, and GloVe are all word-embedding models, i.e., they represent the meaning of words in a high-dimensional space. However, word2vec and fastText both posit a prediction-based mechanism (via neural networks) for acquiring word meanings, whereas GloVe also takes the global structure of language into account during learning. Further, fastText emphasizes the role of learning subword information (while word2vec and GloVe do not), which may represent a more powerful account of how new words are learned and incorporated into memory. Finally, although prior work has found that fastText outperforms word2vec in several semantic tasks [4], there is mixed evidence regarding the performance of GloVe compared to word2vec [1], and their performance in modeling RTs in the free association task remains unknown.

### 2.2 PROCESS MODELS

Spreading activation is a useful metaphor to think about how information spreads over time [5] and proposes that once a word is activated in memory, it gradually activates its semantic neighbors, and these neighbors activate their respective neighbors. A recent study [2] implemented a computational instantiation of the spreading activation model using vectors derived from word2vec and GloVe, but did not investigate free association responses and RTs.

## 3. APPROACH AND UNIQUENESS

This research applies a computational instantiation of the spreading activation mechanism to the free association task and is the first study to examine how RTs in the free association task may be influenced by the underlying structure of semantic information and the dynamic spread of information in the network.

### 3.1. Obtaining Word Vectors

300-dimensional vectors pretrained on a large English Wikipedia corpus (over 5 billion tokens), derived from word2vec, GloVe, and fastText models available from the pymagnitude python package [6] were used for all analyses. We selected 5018 cues from the SWOW database of over 12,000 cues that are also part of the Nelson dataset [7] for all our analyses.

### 3.2. Spreading Activation Model

Based on prior work [2], we implemented the following process model to analyze the predictive power of the embedding models:

1. Three 5018 x 5018-dimensional cosine similarity matrices were constructed from word2vec, fastText, and GloVe vector representations.
2. Negative cosines were set to zero to produce SM, and strength of association between any two words (i, j) was the value of SM (i,j).
3. To implement spread of activation as a discrete-time Markov Chain (MC), diagonals of SM were set to zero, rows were normalized to sum to 1, and converted to a dynamic model, DM = $(2 * SM_{norm} + I_N)/3$. DM took recurrent connections into account using the identity matrix ($I_N$), and other connections between words using $SM_{norm}$.
4. $S_{k\text{-}MC}$ was defined as $(DM)_k$, where k = 1, 2, 3, 4, and 5. For any row i and column j, the value $S_{k\text{-}MC}(i,j)$ represented the amount of activation associated with word $w_j$, at time k, following the initial presentation of word $w_i$.
5. $S_{k\text{-}MC}(i,j)$ and $S_{k\text{-}MC}(j,i)$ were used to estimate the strength of association between $w_i$ and $w_j$, and between $w_j$ and $w_i$, respectively.

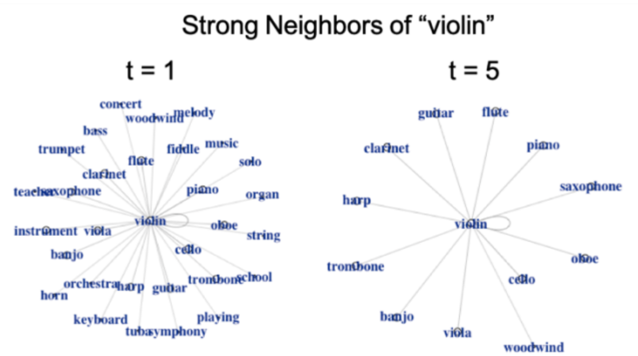


**Figure 1. Semantic neighbors of “violin” with cosine similarity over 0.65 in the 10th decile (i.e., very strong neighbors) in the fastText model at time steps 1 and 5.**

6. $numNeigh_{k,d}(i)$ denoted the number of elements on row i of $S_{k\text{-}MC}$ that had activations (i.e., probabilities) falling into the $d_{th}$ decile of all the activations in $S_{k\text{-}MC}$. This produced the number of active neighbors for any word at each time step. For example, Figure 1 shows the number of active neighbors of “violin” at time steps 1 and 5, i.e., with cosines over 0.65 in the fastText model.
7. Regression analyses were conducted at the word-pair level (cue-target) and predicted (1) response proportions (proportion of participants who responded with an associate given a cue word), and (2) average standardized RTs for word-pair (z-RTs; computed within each participant and then averaged). Models included (a) Base (with word length, frequency, and concreteness of cue, and frequency of associate), (b) Cosine (cosine similarity between cue and associate), (c) T1-5 (included values of $S_{k\text{-}MC}(i,j)$ and $S_{k\text{-}MC}(j,i)$ for k = 1 to 5), and (d) T5+N (included number of top neighbors active at t = 5).

## 4. RESULTS

Figure 2 shows the explained variance (measured by adjusted $R_2$ estimates) for each of the structural and dynamic models at each time step, for response proportions and z-RTs, respectively. As is evident, the spread of activation at each time point significantly improved explained variance for both response proportions as well as z-RTs for the primary response, although the increase in variance was largest from time step 2 to 3 for both response proportions and z-RTs in all models.

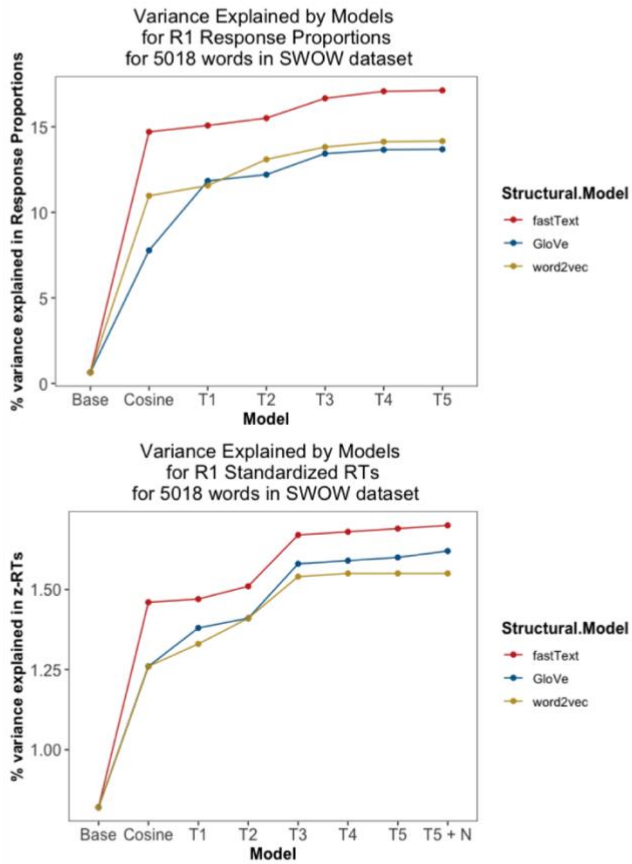


**Figure 2. Percentage of variance explained ($R_2$) for primary (R1) response proportions (top) and z-RTs (bottom)**

Furthermore, fastText performed the best among the different structural models. Therefore, it appears that enriching word representations with subword information (as in fastText) leads to gains in performance, and while cosine similarities are predictive, adding temporal information through spreading activation modeling improves predictive power for free association responses and z-RTs and provides a psychologically intuitive explanation for the success of language models at explaining human performance in the free association task.

## 5. CONTRIBUTIONS

The main contributions of this work are:

- A novel comparison of state-of-the-art embedding models in explaining free association responses and RTs, showing that fastText outperforms other embedding models in this task
- A computational implementation of the spreading activation model [2] combined with word embedding models in a production task, showing how semantic structure and dynamic processes interact to produce human behavior

## 6. FUTURE DIRECTIONS

In this work, we applied a computational method to study the temporal dynamics underlying free association, a phenomenon that is fairly universal and context-dependent. The task can therefore be used to study how free association processes are impacted by different social or political contexts, especially in light of the COVID-19 pandemic. Therefore, examining unconstrained thought processes via free association across different cultural contexts would be an important application and future direction for this research, and using embedding models to study how thoughts propagate within memory networks over time could provide a meaningful lens into cultural transmission of ideas at the global level.

## 7. REFERENCES


[1] Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In Proceedings of the *52nd Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 238-247).

[2] Rotaru, A. S., Vigliocco, G., & Frank, S. L. (2018). Modeling the Structure and Dynamics of Semantic Processing. *Cognitive Science*, 42(8), 2890-2917.

[3] 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.

[4] Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135-146.

[5] Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407.

[6] Patel, A., Sands, A., Callison-Burch, C., & Apidianaki, M. (2018). Magnitude: A fast, efficient universal vector embedding utility package. arXiv preprint arXiv:1810.11190.DOI: https://arxiv.org/pdf/1810.11190.pdf.

[7] Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods*, Instruments, & Computers, 36(3), 402-407