

Investigating Semantic Similarity Effect on Episodic Memory Recall Using Word Embeddings

Burak Büyükyaprak¹, Barbaros Yet¹, Ashlı Kılıç¹

¹ Middle East Technical University

Episodic memory is a type of long-term memory that retrieves personal experiences associated with their context, such as place, time, and associations. Previous episodic memory studies showed that the context of retrieved information may boost or impair the performance of retrieved information. If semantic factors are not controlled, they can be confounders. So, it becomes crucial to study the semantic proximity effect by comparing the study lists that include semantically related or unrelated words. No episodic memory studies that use Turkish words have yet investigated semantic factors by utilizing neural networks. In natural language processing studies, word meanings and relations among words can be successfully represented by learning word vectors in a large text corpus using neural networks. The study aimed to investigate the impact of semantic factors on free recall tasks by creating word lists that include semantically related and unrelated meaningful words obtained through neural networks' word embeddings. In the study, Turkish word representations were obtained by using pre-trained word vectors of fastText and word2vec libraries, and semantically related and unrelated lists were created. A word pool was used that includes 1558 words. Cosine similarity between word vectors was compared to decide if a word should be added to the study list. To create semantically related word lists random words were added to the list regarding the closest word vector. Later, words with the most distant word vectors from the word pool were selected to create a new list. Random words from the word pool were selected once to create semantically unrelated word lists. Later, words from the word pool were added to the list that did not exceed 0.30 cosine similarity, determined by try and error. Therefore, there were ten lists, each containing 12 words for both kinds of lists. The preliminary results are presented to evaluate the quality of the word vectors on the free recall task, Spearman's rank correlation of coefficients conducted between human judgment, and cosine similarity between the word vectors. 278 participants (Mage =31.08; SD = 9.03) were recruited. A significant positive correlation was found between cosine values and human judgment for fastText word embeddings $r(274) = .79, p < 0.001$, and Word2vec word embeddings $r(274) = .66, p < 0.001$. The main experiment was conducted between subject design 2 (word embedding: fastText and word2vec) X 2 (semantic relation: semantically related and unrelated) over ten sessions. A session consisted of three consecutive phases: 1. Participants were instructed to study words on the screen. Participants were required to simultaneously judge whether each presented word was concrete or abstract¹ to reduce rehearsal. 2.

They are given a 60-s arithmetic distractor task to reduce recency effect. 3. The recall task was followed on 90-s. The data is currently being collected from participants. The experiment was designed on PsychoPy. Preliminary results will be presented regarding the impact of semantic associations and temporal organization of words in free recall considering latency and transitions, false recalls, and intrusions.