Project Proposal

## Introduction

* Main premise is based on the Wordnet dataset
* Suggests words in wordnet that share semantic features are also more closely related for people
* Believe semantic features of a concept are important in forming representation in memory – nodes in exemplar theory reference
  + Easier for people to retrieve info about them
  + Concepts that share many semantic features thought to be strongly linked in memory
* In wordnet – concepts of dog and cat share many semantic properties such as mammal, four-legged, pet, animal etc
  + Because of this overlap, easier for people to retrieve concepts from memory

## Testing methodology

* Multiple possible testing methods
* Priming task: participants presented with target word (dog) and quickly presented with related or unrelated word (cat, chair) then asked to indicate if second word is related to first word
  + accuracy: if the participant chooses cat over chair, they are correct. more correct choices = more accurate
  + speed: if the person chooses similar words faster than less similar words. for example if the target word is dog then related and unrelated words are (cow, chair). in this case, cow is less similar than cat so cow should take more time for participant to choose than cat, chair
  + If hypothesis true, can see faster and more accurate responses when the word shares more semantic features with target word in wordnet
  + Otherwise, see no difference in response time
* Reaction time task: speed and accuracy that people can find info about concepts from memory
  + Find some words from wordnet that have different amounts of semantic overlap (like dog and cat vs dog and car), create a set of related words for each pair of words (like bone, hair, pet), also create a set of unrelated words
  + Give participants these words and see how fast they respond to it (test whether they’re related or unrelated)
  + Check if there’s a significant difference in response time for related concepts vs unrelated
  + If hypothesis is correct, there should be a significant difference in response time
* Testing methodology inspired by Auguste et al (2017) based on their experiment on evaluation of word embeddings against cognitive processes with inclusion of reaction time

## Determining results

* Can use t-test to find if the results are significantly different from each other
  + Calc mean and SD of results of similar vs not similar concepts
  + Perform t-test using imported library

## Bibliography

Auguste, J., Rey, A., & Favre, B. (2017). *Evaluation of word embeddings against cognitive processes: Primed reaction times in lexical decision and naming tasks*. 21.

Budanitsky, A., & Hirst, G. (2006). Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics*, *32*(1), 13–47.

Fellbaum, C. (2010). WordNet. In R. Poli, M. Healy, & A. Kameas (Eds.), *Theory and Applications of Ontology: Computer Applications* (pp. 231–243). Springer Netherlands.

Harris, Z. S. (1954). Distributional Structure. *WORD*, *10*(2–3), 146–162.

McDonald, S., & Ramscar, M. (2001). Testing the Distributional Hypothesis: The influence of Context on Judgements of Semantic Similarity. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *23*(23).

Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, *23*(4), 1035–1042.

Patwardhan, S., & Pedersen, T. (2006). *Using WordNet-based Context Vectors to Estimate the Semantic Relatedness of Concepts*. 1–8.