# A Model-Based Analysis of Changes in the Semantic Structure of Free Recall Due to Cognitive Impairment

**Holly A. Westfall (hwestfal@uci.edu)**
Department of Cognitive Sciences, 3151 SSPA UCI
Irvine, CA, 92697, USA

**Michael D. Lee (mdlee@uci.edu)**
Department of Cognitive Sciences, 3151 SSPA UCI
Irvine, CA, 92697, USA

## Abstract

Alzheimer's disease leads to a decline in both episodic and semantic memory. Free recall tasks are commonly used in assessments designed to diagnose and monitor cognitive impairment, but tend to focus only on episodic memory. Our goal is to understand the influence of semantic memory on the sequence of free recall in a clinical data set. We develop a cognitive process model that incorporates the influence of semantic similarity and other stimulus properties on the order of free recall. The model also incorporates a decision process based on the Luce choice rule, allowing for different levels of response determinism. We apply the model to a real-world data set including free recall data from 2392 Alzheimer's patients and their caregivers. We find that semantic similarity between items strongly influences the order of free recall, regardless of impairment. We also observe a trend for response determinism to decrease as impairment increases.

**Keywords:** semantic memory; episodic memory; free recall; odd-one-out task; Alzheimer's disease

## Introduction

Alzheimer's disease causes changes in memory that are distinct from normal aging. In cognitively healthy adults, memory for past events (episodic memory) tends to decline over time, but memory for facts (semantic memory) typically remains intact (see Balota et al., 2000, for a review). Memory deficits over and above those seen in healthy aging can be a sign of cognitive impairment (Nebes & Brady, 1990). Patients with Alzheimer's disease display deficits in both episodic and semantic memory early in the disease, with deficits in memory increasing with the severity of cognitive impairment (Mortamais et al., 2017).

A commonly used task in memory assessment is the free recall task, which is widely used to study memory and decision processes. In this task, a participant is presented with a list of items and then – immediately or after some delay – must recall those items, in any order. There are several common ways to analyze free recall data. Techniques such as serial position analysis, in which effects of primacy and recency are typically observed (Murdock, 1962), rely on the order of items presented at study. Conditional response probabilities and lag-recency effects (Howard & Kahana, 1999) are concerned with the temporal relationship of successively recalled items at test. Simple recall accuracy is a quick and easy way of analyzing memory performance, but it is not sensitive to any information about the presentation order of items or the order of recall output.

Often, in a recall task, stimuli are controlled on various dimensions, such as word frequency, word length, age of acquisition, emotional valence, or semantic similarity. These kinds of controls are implemented because semantic memory can influence episodic memory. For example, people tend to recall words in semantically-related clusters, both between categories (Bousfield, 1953) and within categories (Romney et al., 1993). Because of this, stimuli in free recall tasks tend to be controlled for their semantic relationships, to prevent semantic associations being used in recall. These controls, however, limit the ability of free recall behavior to inform our understanding of a basic feature of human memory, which is to organize stimuli and represent meaning. Understanding the semantic relationships between successively recalled items can provide valuable insight into memory and decision processes, including how semantic memory is affected by impairment in Alzheimer's patients (Ribeiro et al., 2007).

The goal of this paper is to model the sequence of free recall in a clinical data set in an effort to understand how semantic similarity and other features of stimuli guide the recall process. In the next section, we describe the clinical data set and an initial analysis of first-order transition probabilities. We then describe a cognitive model of free recall output based on the similarity to other recalled words, as well as several other properties of the words themselves. We find that people do not tend to use different cues as they become more impaired. Instead, free recall output tends to become more random as impairment increases, as measured by a cognitive parameter corresponding to response determinism. Finally we discuss the implications and limitations of our findings.

## Data

The data were collected as part of a routine assessment of Alzheimer's patients and their caregivers at a clinic specializing in neurodegenerative disorders. All participants completed the Mild Cognitive Impairment Screen (MCIS: Shankle et al., 2009), which is used to help diagnose and monitor cognitive impairment. This screen includes an odd-one-out comparison of animal names task and an unexpected free recall task of those animal names. The odd-one-out comparison task draws stimuli from a pool of 21 animal names: antelope, beaver, camel, cat, chimpanzee, chipmunk, cow, deer, dog, elephant, giraffe, goat, gorilla, horse, lion, monkey, rabbit, rat, sheep, tiger, and zebra. For each triad of

Table 1: Identifying characteristics for each FAST stage, the number of participants in each stage, and descriptive statistics for the number of words correctly recalled.

| Stage | Description | n | M | SD |
|---|---|---|---|---|
| 1 & 2 | no deficit or subjective deficit | 518 | 6.8 | 1.6 |
| 3 | objective deficit in complex tasks | 782 | 5.6 | 1.9 |
| 4 | mild dementia evident in IADLs | 770 | 4.1 | 2.0 |
| 5 | moderate dementia | 152 | 3.5 | 2.0 |
| 6 | moderately severe dementia | 170 | 3.0 | 1.8 |

animal names, the participant must choose which animal is least like the other two. For example, if presented with the words "cow", "elephant", and "giraffe", a person might choose "cow" as the odd one out. The clinician does not offer any feedback after each choice, as there is no correct answer for this task. In accordance with a $\lambda$-2 balanced, incomplete block design (Burton & Nerlove, 1976), nine animal names are drawn from the pool for each participant, and each of the selected animals is presented verbally in a triad with every other animal over the course of 12 trials. After a delay, in which participants complete other unrelated tasks, there is an unexpected free recall task of these animal names. The instructions are to try to recall as many of the animal names as possible, in any order.

The data set includes assessments from 2392 participants (52% female, age range 16–101 years, mean age 74 years). At the time of assessment, all participants were also classified using the Functional Assessment Staging Test (FAST: Reisberg, 1988). The FAST assessment is an evaluation of a person's ability to perform Instrumental Activities of Daily Living (IADLs: Lawton & Brody, 1969), such as cooking, cleaning, and managing finances, as well as Activities of Daily living (ADLs: Katz et al., 1963), such as dressing, bathing, and grooming. Participants in FAST stages 1 and 2 have either no functional deficit or only a subjective deficit and are considered to be cognitively healthy for the purposes of this analysis. Those in stage 3 have mild cognitive impairment and are beginning to show an objective deficit in accomplishing more complex tasks. Participants in stages 4, 5, and 6, have been diagnosed with mild, moderate, and moderately severe dementia, respectively. The FAST assessment is made independently of the MCIS, and so it provides a way to group participants by impairment, in order to study changes in free recall. A summary of the number of participants grouped into each FAST stage is presented in Table 1.

In an initial analysis of the data, we wanted to determine whether the semantic similarity of the animal names influenced the order of free recall output, and if so, whether this influence varied with cognitive impairment. First we inferred animal similarity data from the odd-one-out comparison task.[1] For each pair of animal names, we counted the number of times neither animal was chosen as the odd one out in a triad and divided that by the total number of times that pair was presented in a triad. In effect, this means that the similarities of pairs of animals increase to the extent neither is chosen as the odd one out when they are both present in a triad. Then we calculated first-order transition probabilities to measure the probability of the next recalled word given the most recently recalled word, similar to semantic conditional response probabilities (Howard & Kahana, 2002).

In Figure 1, for each stage, all 21 animals are presented on the circumference of a circle that arranges them by similarity. Transition probabilities are represented by arrows of varying width. Higher transition probabilities are represented by thicker arrows and lower transition probabilities are represented by thinner arrows. Transitions occurring around the edge of the circle are indicative of similarity-based search, while transitions through the middle of the circle indicate some other process, inconsistent with similarity. In stages 1 & 2, most of the transitions occur around the edge of the circle, between animals that are semantically similar to each other. Transitions between similar animals such as "gorilla", "chimpanzee", and "monkey" are common, while transitions between dissimilar animals, such as "camel" and "dog", are much less common. In stages 3 and 4, many of the transitions occur around the edge of the circle, between semantically similar animals, but there is a relatively greater number of transitions that cross through the center of the circle between dissimilar animals. In stages 5 and 6, there appear to be as many transitions between similar animals as dissimilar animals, with a great many transitions crossing through the center of the circle. From these plots, it is clear that as FAST stage increases, transitions are more and more likely to cross through the middle of the circles. In other words, as cognitive impairment increases, transitions become less consistent with similarity of the animal names.

## Model

A description of the model follows and is broken into two parts: a regression function, allowing for the influence of between- and within- stimuli factors on memory, and a decision-making process that produces choice probabilities. We applied the model to each FAST stage separately, rather than attempting to estimate a linear or otherwise monotonic progression across FAST stages. While our purpose was to measure change across FAST stage, we did not want to incorporate such a strong assumption into the model itself.

The model was written in R (R Core Team, 2018) and implemented in JAGS (Plummer, 2003) via the rjags package

[1] There are many ways to calculate similarity, including free association measures (De Deyne et al., 2019; Nelson et al., 2004), latent semantic analysis (LSA: Landauer & Dumais, 1997), and vector cosine similarity via an algorithm such as word2vec (Mikolov et al., 2013). We measure animal similarity from the odd-one-out comparison task following Romney et al. (1993) because the odd-one-out data and the free recall data were collected from the same people.

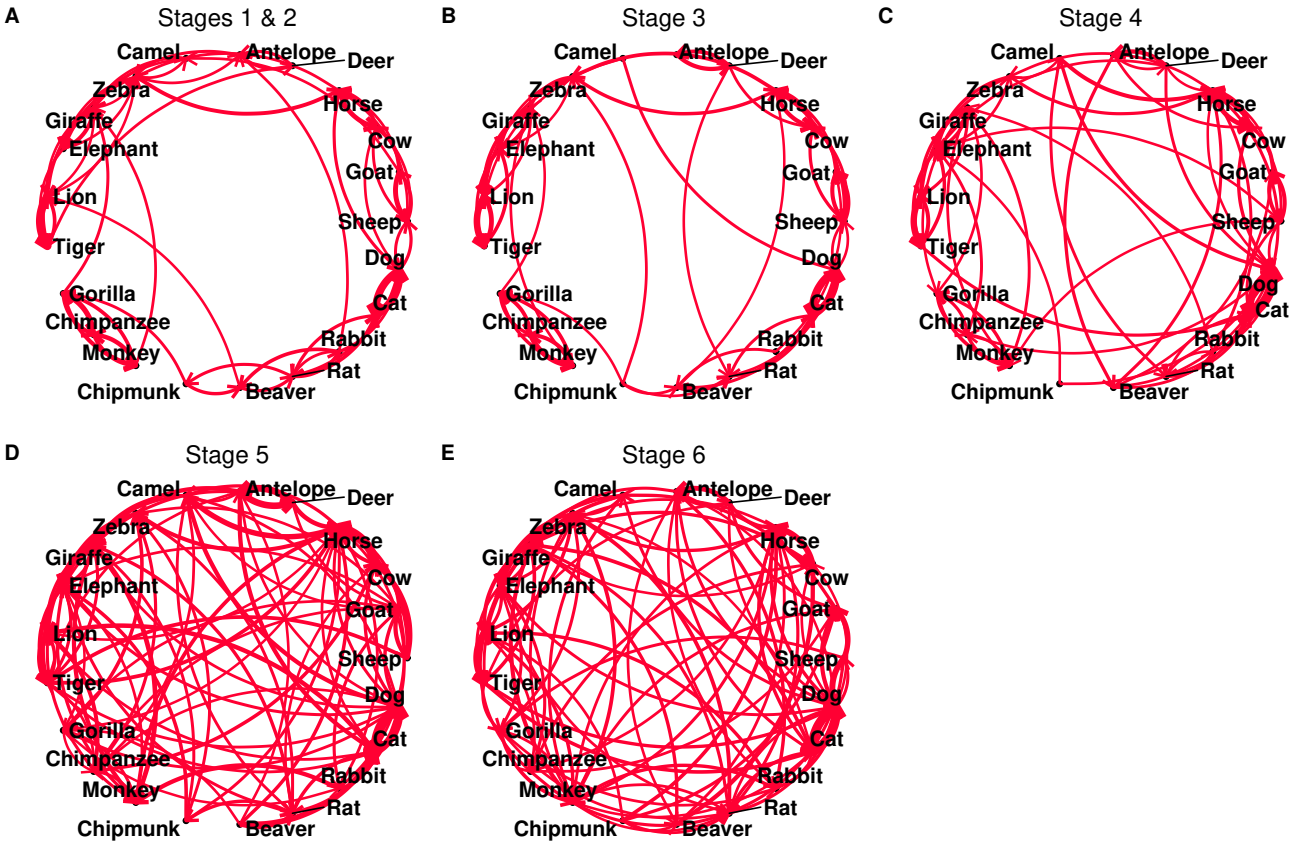


Figure 1: First-order transition probabilities for the animal free recall task. Animals are arranged around the circumference of each circle in terms of similarity. Transitions are represented as red arrows, where higher transition probabilities are represented by thicker arrows and lower transition probabilities are represented by thinner arrows. As impairment increases, transitions become less consistent with animal similarity.

(Plummer, 2016). JAGS provides a high-level scripting language for implementing probabilistic cognitive models that allows for computational Bayesian analysis using Markov-chain Monte Carlo sampling methods (Lee & Wagenmakers, 2014). The results are based on three chains of 10,000 posterior samples collected after 1000 discarded burn-in samples, with a thinning factor of 10. We assessed convergence of chains by visual inspection and through the $\hat{R}$ statistic (Brooks & Gelman, 1998).

## Regression function

The regression function contains weights of theoretical interest $\alpha$, which express the influence of various psychological features on recall probabilities. Specifically, we assume that the memory strength $Q$ allowing for the recall of animal $j$ on trial $k$ given that animal $i$ has just been recalled on trial $k-1$ (i.e., $y_{k-1} = i$) is a function of both the similarity of animal $j$ to animal $i$, and specific features of word $j$:

$$Q_k(j|y_{k-1} = i) = \alpha_1 \times \text{similarity}_{ij} + \alpha_2 \times \text{oddness}_j \quad (1)$$
$$+ \alpha_3 \times \text{frequency}_j - \alpha_4 \times \text{length}_j - \alpha_5 \times \text{aoa}_j - \alpha_6 \times \text{valence}_j.$$

For each animal name, we obtained the word frequency, word length, age of acquisition, and emotional valence from The English Lexicon Project (Balota et al., 2007). Word frequency corresponds to the log-transformed Hyperspace Analogue to Language (HAL: Lund & Burgess, 1996) frequency norms. Word length is defined as the number of syllables. We also include oddness, which is the number of times a participant chose an animal name in the odd-one-out comparison task, as a word-level feature. All of the predictors were rescaled to be in the range of 0 to 1.

We assume each FAST stage has its own regression weights, and they are given a Dirichlet prior. This choice of prior ensures that the weights sum to one and allows us to interpret the weights as the relative importance of each predictor on memory:

$$\boldsymbol{\alpha} \sim \text{Dirichlet}(1,1,1,1,1,1). \quad (2)$$

## Luce Choice Rule

Since there is a discrete number of possible animal names, the free recall task can be thought of as a multinomial choice task. We use the Luce choice rule (Luce, 1959) extended to

incorporate response determinism, which allows us to assign a probability to each of these response options. According to the Luce choice rule, choice probabilities $S$ are defined as:

$$S(j) = \frac{\exp(\gamma Q(j))}{\sum_m \exp(\gamma Q(m))}, \tag{3}$$

where $j$ refers to one of the recalled animal names, the summation in the denominator is over all recalled animal names, and the function $Q$ is defined in Equation 1 (and the conditioning on the previous observation $y_{k-1}$ is now implicit). Finally, the response determinism parameter $\gamma$ determines the degree to which a participant chooses a response consistent with the regression function $Q$. If $\gamma = 0$, $Q$ is ignored, and the probability of choosing option $j$ becomes $\frac{1}{m}$. If $\gamma = 1$, decisions become consistent with probability matching. As $\gamma$ increases, the decision becomes more deterministic based on $Q$, and a participant will eventually always pick the option with the highest $Q$. In this way, the determinism parameter can be interpreted as the consistency of decisions. Following Lee et al. (2016), we assume each FAST stage has its own response determinism, and assume a gamma prior:

$$\gamma \sim \text{gamma}\big(2,1\big), \tag{4}$$

which has a mode corresponding to probability matching, but allows for higher and lower values.

**Censoring repeated recall.** For the purposes of this analysis, we removed extra-list intrusions, but it is still possible for the participant to recall the same word repeatedly. Because this type of task error could be related to cognitive impairment, we extend our model to include a censoring component that captures the ability to inhibit repeated recalls.

The probability of choosing a response option on a particular trial is determined by the Luce choice rule and a censoring index $\delta$, which is a binary indicator that takes the value 1 for words that have not been previously recalled. This index can be pre-calculated and treated as observed data.

Whether a participant is in a state of censoring on a trial is determined by parameter $z$. If $z = 0$, the participant *is not* in a censoring state, and the probability of recalling alternative $j$ is determined by the Luce choice rule. If $z = 1$, then the participant *is* in a censoring state, and the probability of recalling alternative $j$ is equal to $\delta$ times the Luce choice rule:

$$P_k(y_k = j) = \begin{cases} S(j), & \text{if } z_k = 0 \\ \delta_j S(j), & \text{if } z_k = 1. \end{cases} \tag{5}$$

The parameter $z$ has a Bernoulli prior with hyperparameter $\phi$:

$$z_k \sim \text{Bernoulli}(\phi). \tag{6}$$

In our application of the model $\phi$ is a person-specific parameter that captures the individual's ability to censor previously recalled words. $\phi$ can range from 0 (no censoring) to 1 (perfect censoring), and is given a standard uniform prior:

$$\phi \sim \text{uniform}\big(0,1\big). \tag{7}$$

### First word recall

For first word recall our regression equation changes slightly, because recall of the first word depends only on the properties of the individual word, and not on the similarity to a previously recalled item. Additionally, there is no need for a censoring component for first word recall. To distinguish first word recall from subsequent word recall, we denote the regression weights with $\beta$ (instead of $\alpha$) and the response determinism parameter with $\kappa$ (instead of $\gamma$). The regression function governing first word recall is now:

$$\begin{aligned} Q_{k=1}(j) = {} & \beta_1 \times \text{oddness}_j + \beta_2 \times \text{frequency}_j \\ & - \beta_3 \times \text{length}_j - \beta_4 \times \text{aoa}_j - \beta_5 \times \text{valence}_j. \end{aligned} \tag{8}$$

As before, the regression weights are given a Dirichlet prior:

$$\boldsymbol{\beta} \sim \text{Dirichlet}(1,1,1,1). \tag{9}$$

The probability of choice in Equation 3 becomes:

$$S(j) = \frac{\exp(\kappa Q(j))}{\sum_m \exp(\kappa Q(m))}, \tag{10}$$

where $\kappa$ has the gamma prior:

$$\kappa \sim \text{gamma}\big(2,1\big). \tag{11}$$

## Results

We have a clear expectation for response determinism. As impairment increases, free recall output should become less consistent with the memory of the items. In other words, response determinism should decrease as FAST stage increases. However, we are also interested in the relative influence of each of the regression predictors on the structure of free recall, and whether they change across impairment.

### Model descriptive adequacy

We quantified model fit by creating a confusion matrix for each FAST stage that compares the true word recall of the participants to the model-described word recall. From these matrices we were able to calculate the overall descriptive accuracy of the model, which was between 40% and 57% for each stage. A model that predicted outcomes completely randomly would have an accuracy of only $\frac{1}{21}$, or 5%. While far from perfect, this suggests that the model provides a reasonably accurate description of the data, and performs an order of magnitude better than a random chance model.

### Regression weights

The posterior distributions for the regression weights $\alpha$ of each of the predictors are presented in Figure 2. The regression weights for similarity are very high overall, meaning that the similarity between items is particularly important in the sequence of free recall. Oddness and emotional valence have relatively less influence on recall order compared to similarity. The weights for the other item-dependent properties,

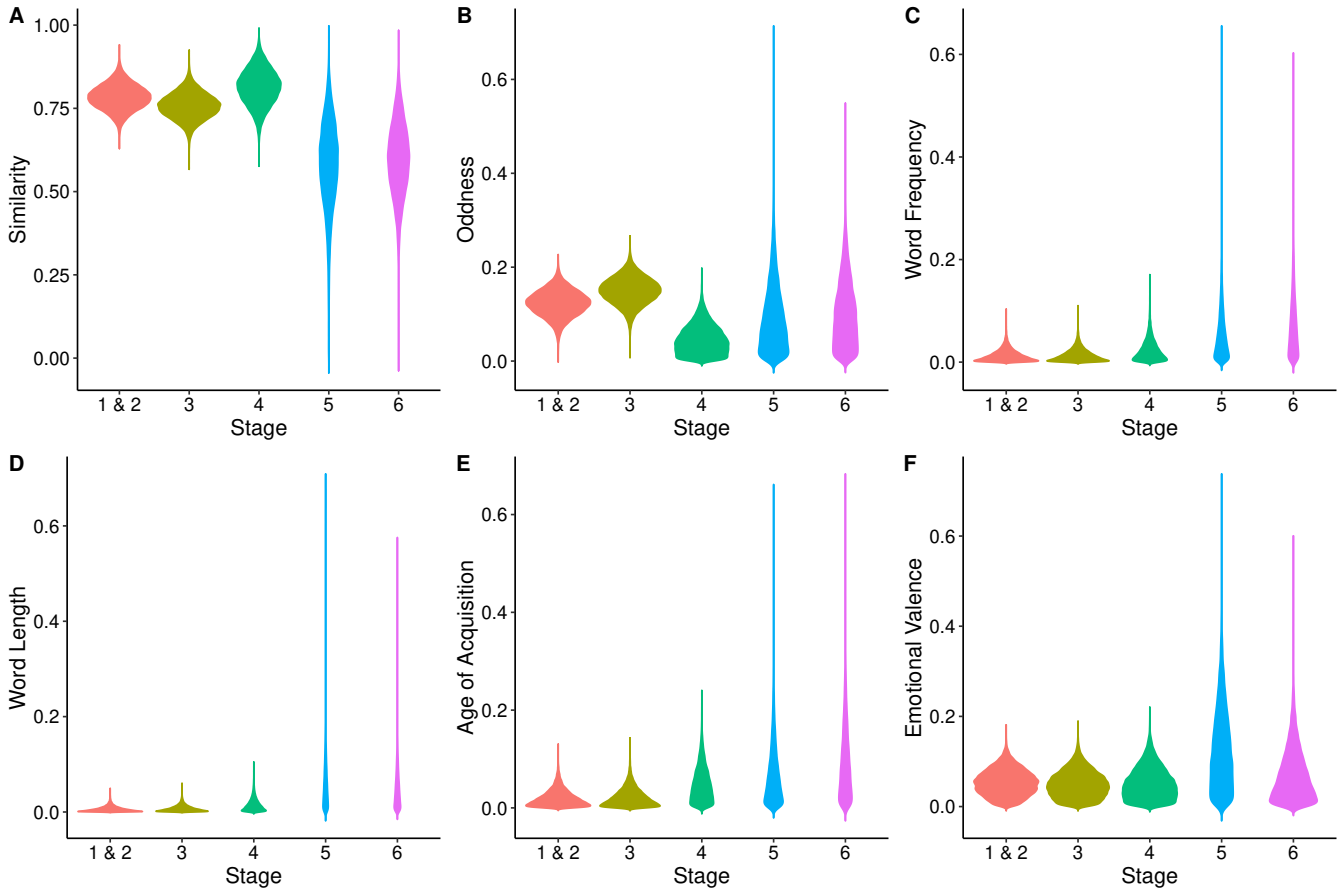


Figure 2: Posterior distributions of regression weights $\alpha$ for each FAST stage. The panels represent (A) similarity between items, (B) probability of odd-one-out selection, (C) word frequency, (D) word length, (E) age of acquisition, and (F) emotional valence.

word frequency, word length, and age of acquisition, are all very close to zero.

It appears that the regression weights are fairly stable across stages. To quantify the stability of the weights across FAST stage, we calculated Bayes factors between adjacent stages for each of the regression weights. Our results are presented in Table 2 as Bayes factors in favor of the null. Overall, the regression weights do not tend to shift across FAST stage. In other words, people do not tend to use different cues in the free recall task as they become more impaired.

## Response determinism

The posterior distributions for response determinism $\gamma$ are presented in Figure 3. There seems to be a visual trend wherein determinism tends to decrease as FAST stage increases. Again, we computed Bayes factors to quantify any differences between adjacent FAST stages. In this case, the Bayes factors in favor of the null were 2.8, 4.5, 1.5, and 1.5 for comparisons between stages 1 & 2 vs. stage 3, stage 3 vs. stage 4, stage 4 vs. stage 5, and stage 5 vs. stage 6, respectively. Response determinism is trending in a way that would suggest free recall output becomes less consistent with the model and more random as impairment increases. However, the Bayes factor magnitudes represent suggestive rather than clear evidence of differences between adjacent stages.

## First-word recall

While subsequent word recall was well-described by the regression model and Luce choice rule, first-word recall was not. The posterior mean regression weights for oddness ranged between 0.41 and 0.80 across impairment, while the mean regression weights for emotional valence ranged between 0.09 and 0.26. The other predictors – word frequency, word length, and age of acquisition – had mean regression weights all very close to zero for all FAST stages. However, interpretation of these regression weights is complicated by the fact that the mean posterior response determinism is either very close to one, in the case of stages 1, 2, and 3, or less than one, for FAST stages greater than 3. This observation means that many participants are responding in a way that is more random than probability matching. It is possible that we failed to incorporate relevant predictors for first-word recall into the model, or that our modeling assumptions are otherwise inappropriate for first-word recall. Accounting for

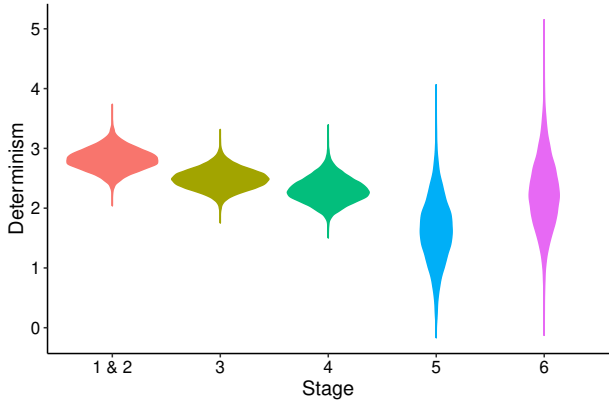


Figure 3: Posterior distributions of response determinism γ for each FAST stage. Response determinism tends to decrease as impairment increases.

first-word recall requires more attention in further model development.

## Discussion

We created a cognitive process model to try to understand the sequence of free recall output in a clinical data set. The model has two main components: a regression-type equation that describes the influence of within- and between-item factors, and a decision process based on the Luce choice rule. We found that the similarity between words has a very high influence on word choice for all stages, while oddness and emotional valence have a smaller influence. Word frequency, word length, and age of acquisition seem to matter very little in terms of word choice for this particular set of stimuli. The fact that these word-specific characteristics had little influence on the order of recall is not surprising in this particular context. These animal names were chosen for this task to be similar on several dimensions. However, in all cases, the weights of each of these between- and within-item factors was consistent across all FAST stages. In other words, participants did not appear to use different cues for word choice, regardless of impairment.

Response determinism has a trend such that determinism decreases as FAST stage increases. A decrease in response determinism means that decisions become less consistent with the regression model and more random. However, Bayes factor comparisons between successive stages revealed that any differences were not large enough to make strong claims. Our use of a response determinism parameter is similar to the model described in Lee et al. (2016). We assume that all participants have the same underlying semantic representation, but the ability to access that information decreases as impairment increases (Nebes & Brady, 1990). In other words, cognitive impairment limits the ability to access stored semantic information in memory. This is essentially the same conclusion about the impacts of impairment reached by Westfall & Lee (in press) in their model-based analysis of the odd-one-

Table 2: Bayes factors for the comparison of each regression weight across adjacent FAST stages. All Bayes factors are presented in favor of the null.

| | 1&2 vs. 3 | 3 vs. 4 | 4 vs. 5 | 5 vs. 6 |
|---|---|---|---|---|
| Similarity | 2.4 | 1.5 | 0.3 | 0.9 |
| Oddness | 3.0 | 0.2 | 2.1 | 1.6 |
| Frequency | 12.6 | 8.4 | 4.0 | 2.4 |
| Length | 25.9 | 13.8 | 3.8 | 2.7 |
| AOA | 8.0 | 4.3 | 2.6 | 2.0 |
| Valence | 3.6 | 3.3 | 1.5 | 1.5 |

choice behavior that we used to determine semantic similarity.

A major assumption of this model is that we know what items will be recalled. The model is only concerned with the order of free recall output given the words that were recalled successfully. A complete account of the sequence of free recall should be able to predict the words that will be recalled, given the memory stimuli, as well as providing a better account of first-word recall.

Nevertheless, the modeling results advance our understanding of the factors that influence the free recall of within-category items. In particular, the use of a cognitive model allows us to identify patterns and relationships not observable in standard data analysis. Measuring latent psychological parameters potentially provides a more precise measurement of semantic clustering than other more common behavioral methods, such as the California Verbal Learning Test (CVLT-II: Delis et al., 2000). The model allows us access to information that would otherwise be lost in a typical account of free recall accuracy, and helps us understand recall order when information on study order is unavailable or ambiguous. Most importantly, our modeling provides an insight into how semantic memory drives free recall, and how this interaction changes with memory impairment.

## Acknowledgments

This work was financially supported by a gift from Medical Care Corporation. We thank Rod Shankle, Jason Bock, Dennis Fortier, and Junko Hara, and members of the Bayesian Cognitive Modeling lab at UC Irvine, for helpful discussions.

## References

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 395–410). Oxford University Press, Inc.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology*, *49*, 229–240.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434–455.

Burton, M. L., & Nerlove, S. B. (1976). Balanced designs for triads tests: Two examples from English. *Social Science Research*, *5*, 247–267.

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.

Delis, D. C., Kramer, J. H., & Kaplan, E. (2000). *California Verbal Learning Test* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 923–941.

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*(1), 85–98.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychosocial function. *Journal of the American Medical Association*, *185*, 914–919.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*, 179–186.

Lee, M. D., Abramyan, M., & Shankle, W. R. (2016). New methods, measures, and models for analyzing memory impairment using triadic comparisons. *Behavior Research Methods*, *48*, 1492–1507.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Luce, R. D. (1959). *Individual choice behavior*. John Wiley & Sons.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*, 203–208.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., ... Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia*, *13*, 468–492.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488.

Nebes, R. D., & Brady, C. B. (1990). Preserved organization of semantic attributes in Alzheimer's disease. *Psychology and Aging*, *5*, 574–579.

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.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. Vienna, Austria.

Plummer, M. (2016). rjags: Bayesian graphical models using MCMC [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=rjags` (R package version 4-6)

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Reisberg, B. (1988). Functional assessment staging (FAST). *Psychopharmacology Bulletin*, *24*, 653–659.

Ribeiro, F., Guerreiro, M., & De Mendonça, A. (2007). Verbal learning and memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *29*(2), 187–197.

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, *4*, 28–34.

Shankle, W. R., Mangrola, T., Chan, T., & Hara, J. (2009). Development and validation of the Memory Performance Index: Reducing measurement error in recall tests. *Alzheimer's & Dementia*, *5*, 295–306.

Westfall, H. A., & Lee, M. D. (in press). A model-based analysis of the impairment of semantic memory. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-020-01875-9