UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Using "Semantic Scent" to Predict Item-Specific Clustering and Switching Patterns in Memory Search

Permalink

https://escholarship.org/uc/item/67m4g3d9

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Zhang, Larry
Jones, Michael N.

Publication Date

2022

Peer reviewed

Using "Semantic Scent" to Predict Item-Specific Clustering and Switching Patterns in Memory Search

Larry Zhang (larzhang@iu.edu)
Department of Informatics
Department of Intelligent Systems Engineering
Indiana University Bloomington

Michael N. Jones (jonesmn@indiana.edu)
Department of Psychological and Brain Sciences
Indiana University Bloomington

Abstract

Elucidating the mechanisms that underlie clustering and switching behavior is essential to understanding semantic memory search and retrieval. Hills, Jones, and Todd (2012) proposed a model of semantic foraging based on the observation that statistical signatures in memory search resemble optimal foraging in animal behavior. However, the original model was postdictive in explaining when a switch would occur, as opposed to predictive, and was agnostic as to the cues used by humans to make a decision to switch from local to global information. In this paper, we proposed a switching mechanism, *Semantic Scent*, as a predictive model underlying such behavior. Semantic Scent extends optimal foraging theory, reproducing the same switch behavior observed animal foraging behavior in memory search. We evaluated Semantic Scent against competing models including *Random Walk* and *Fixed Count* to determine its effectiveness in classifying switches made in fluency tasks. A quantitative model comparison between the switch models demonstrated Semantic Scent's superior performance in fitting human data. These results provide further evidence of the importance of optimal foraging theory to semantic memory search.

Keywords: Semantic Memory; memory search; clustering and switching

Introduction

Understanding the mechanisms humans use to search for and retrieve information from semantic memory has been a primary goal of cognitive science since the field's inception. One of the most commonly used tasks to study semantic search is the semantic fluency task (SFT). In the SFT, the participant is presented with a category label (e.g., "animals") and is asked to produce as many exemplars of the category as possible within a fixed amount of time (e.g., dog, cat, spider, ant, ...). SFT is commonly used in experimental psychology (Raaijmakers & Shiffrin, 1981; Romney, Brewer, & Batchelder, 1993), but is also widely used in neuropsychological batteries.

Responses in SFT typically occur in temporal clusters of related items (e.g., {farm animals} {pets} {fish}). Counts of clustering and switching within a sequence of items produced in a trial are known to be particularly sensitive to clinical group diagnoses in Alzheimer's and Parkinson's Disease (Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998) as well as Schizophrenia (Lundin et al., 2020) among other clinical populations. Clustering and switching is typically coded by hand using the classification scheme proposed by Troyer et al. (1998). While clusters in semantic memory are no doubt more complex than the simple scheme proposed by Troyer et al. (1998; see (Hills, Todd, & Jones, 2015) for a review), their simple method has been very widely applied in both experimental and clinical settings and has impressive predictive validity.

Hills, Jones, & Todd (2012) made the observation that the temporal pattern of items produced in SFT exhibited statistical signatures that are characteristic of animals foraging for food in physical space (optimal foraging theory: (Charnov, 1976)), suggesting that our memory search mechanisms may have been exapted from primitive mechanisms that evolved to search for food resources in the physical environment (Hills, Jones, & Todd, 2012). They tested a series of search models on a semantic space generated by the BEAGLE model of semantic memory (Jones & Mewhort, 2007). The specific search model that best explained the human data was a dynamic two-cue model that used semantic similarity locally to generate items until no other proximal item was found, and then switched to a global frequency cue to select the next item (and search by local similarity resumed). The fact that the local-global switch model produced the best fit to the human data was theoretically significant for two reasons: 1) it produced patterns of optimal foraging, and 2) the process it used mirrors the best accounts of how animals make exploration-exploitation decisions when foraging for food in physical environments. Just as a hummingbird must decide when to give up on a local patch of flowers and accept the costs that accompany the search for a new unknown patch, humans show the same pattern in memory search when deciding when to give up on the farm animals and search for a new resource-rich semantic patch to exploit.

However, it is important to note that the optimal foraging model presented by Hills, Jones & Todd (2012) did not predict *when* a switch would occur. Rather the model relied on either Troyer hand-coded norms of when a switch occurred or used a similarity-drop heuristic which classified a switch as having occurred if semantic similarity dropped between successive items. For example, if S(A, B) represents the similarity between retrieved words A and B, then a switch following B is identified in a series of retrievals A, B, C, D if S(A, B) > S(B, C) and S(B,C) < S(C, D). The foraging model sug-
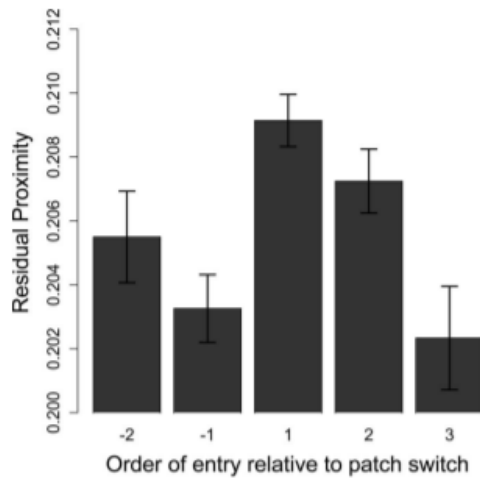
Figure 1: Figure 3 reproduced from (Hills et al., 2012) demonstrating the reduction in residual proximity before switching and spiking of residual proximity after switching

gested that humans were abandoning local similarity cues and switching to a global frequency cue when a switch occurred (indicated either by the Troyer or the similarity-drop method) and the switching between these cues best explained the sequence of items produced. This is analogous to an animal balancing exploration-exploitation by foraging locally on a bush for berries based on visual cues, but then abandoning that cue with diminishing local returns and switching to a scent based cue to locate a new patch ripe with berries to then revert to visual cues for local search again.

A primary shortcoming of the optimal foraging model of Hills et al. (2012) is that it is postdictive, not predictive, of a switch from a semantic cluster to another. It relies on either Troyer hand coded norms or on noting that similarity dropped and thus a switch had occurred. As a major goal of cognitive modeling, models should seek to explain behavior as it happens (*predictive*), rather than be explaining behavior as a post-hoc judgement (*postdictive*). This shortcoming was partially addressed in the original paper by noting that it appeared that humans were making decisions to abandon local semantic foraging within a cluster and switch to a global frequency cue to find a new cluster when the sequence of items had reached a point where there were few semantically proximal items to the one just produced. Parallel to the diminishing returns from a local patch of berries, Hills et al. (2012; Fig 3) showed that the residual semantic proximity decreased as items were produced in a cluster, but then spiked again when a switch occurred and a new semantically rich patch of items was located. Residual proximity was simply quantified as the mean similarity to all yet-to-be-produced items in memory, and indicates how semantically "urban" or "rural" the neighborhood is around the item just produced. Hills et al. noted that this pattern indicates that humans are making a decision

to switch cues from local (similarity) to global (frequency) when there are insufficient proximal items in semantic memory to the one just produced, much like the animal that abandoned visual cues and switched back to scent cues to find a new patch of berries.

Although this is indicative of a mechanism, the model was descriptive—no predictive model was created or tested. In this paper, we formalize a mechanism to predict when a switch will occur in a sequence of items and test the model's predictions of switches against competing models. We refer to the mechanism as Semantic Scent, borrowing from scent-based models switching models in animal ecology and from Pirolli and Card's construct of information scent in web search (Pirolli & Card, 1999). The semantic scent after producing an item relates to the number and proximity of yet-to-be produced items in mental space, and this scent is used in a simple decision whether to continuing searching locally based on semantic information, or to abandon local search and switch to a frequency cue to "fly off" to another patch in semantic memory and revert to local semantic search. We compare the Semantic Scent model's predictions of when switches occurred in a large set of SFT data to the competing Random Walk model of (Abbott, Austerweil, & Griffiths, 2015) and to a simple count-based model that is insensitive to the semantics of the items produced but is a standard baseline heuristic model in animal foraging.
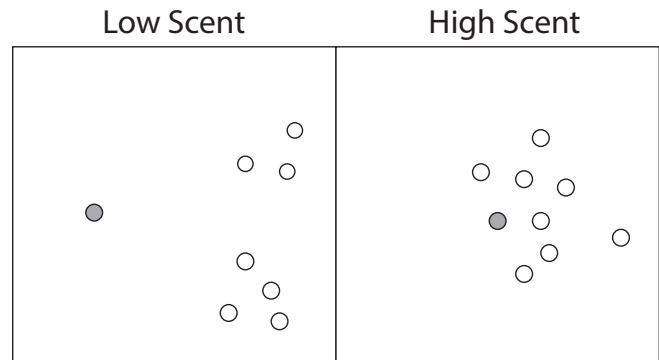

Figure 2: Semantic Scent drops as resources become more scarce in a patch. Scent spikes again after transitioning to a new patch rich with resources

Memory Search Switching Models

Semantic Scent Model

The proposed Semantic Scent (SS) model is a predictive model of clustering and switching behavior in semantic memory search. Inspired by Pirolli's information scent model, the semantic scent model operates as a method of quantifying the content of proximal cues to develop a value system in semantic search (Chi, Pirolli, Chen, & Pitkow, 2001). This follows similar phenomenology to information foraging in web searching behavior. The scent of the local area drops as resources become scarce in a local patch. Scent spikes again

after transitioning to a new patch rich with resources, as seen in **Figure 2**.

For the purpose of analysis, conceptual items named in the semantic fluency task are conceived as embedded items in a distributional semantic space. Items in the distributional semantic space form variably dense patches of available semantic resources. The semantic scent model predicts the likelihood of switching from patch to patch based on the item just produced and the proximity to yet-to-be-produced items.

Semantic scent is measured as a power function, based on the residual proximity of items to the just produced item in the distributional semantic space, raised to a standard sensitivity parameter based on Shepard's Universal Law of generalization (Shepard, 1987).

$$P(Switch|X,N) = \frac{1}{1 + \sum_{i=1}^{N} cos(X, Y_i)^\lambda} \quad (1)$$

$$P(ContinueLocal|X,N) = 1 - P(Switch) \quad (2)$$

Y corresponds to the vectors of N most similar items to X determined by the distributional semantic space utilized in the model. Y_i corresponds to the i-th most similar item to the current item X in the semantic space. The model is optimized on two dimensions including the neighborhood parameter (N) and the sensitivity parameter (λ). The neighborhood parameter corresponds to the number of items included in the residual proximity calculation, determined by the closeness in similarity to the current item.

Random Walk Model

The Random Walk (RW) model is based on the generative process of random traversal on a semantic network (Abbott, Austerweil, & Griffiths, 2012). Generated fluency lists are based on stochastic jumps from node to node in the semantic network by following the edge connections, also known as a random walk. Similar to the Semantic Scent model, the Random Walk model derives its' inspiration from information retrieval research, notably the PageRank algorithm (Page, Brin, Motwani, & Winograd, 1998). The original application of PageRank was to model the information retrieval process on the World Wide Web, which was conceived as a network of web pages. Each node encoded a unique webpage, and each directed edge corresponded to hyperlinks to other webpages. The PageRank algorithm derives the largest eigenvector of a transition probability matrix of the network, estimating "the limiting distribution of a random walk on the web graph" (Page et al., 1998). The resulting matrix produced by PageRank reflects the probability of visiting on any unique webpage in the network.

The Random Walk model utilized in the original Abbott, Austerweil, and Griffiths paper develops a generative model with two simultaneous components. The first component is

based on the transition probability of the current item to the next item in the fluency list based on the network transition probability. The second component is based on a jumping cue, analogous to the stochastic restart of the random walk from the original PageRank model. We utilize the original jumping cue from Abbott, Austerweil, and Griffiths (2012), where the random walk model jumps back to the item "animal" in the network, in order to subsequently transition to the next item in the fluency list.

$$P(X_{n+1}|Q1, Q2, N) = \rho \cdot P(X_{n+1}|Q_2) + (1-\rho) \cdot P(X_{n+1}|Q_1(X_{n+1}, X_n)) \quad (3)$$

$Q1$ corresponds to the to the first component based on the transition probability between the current item and next item in the fluency list. $Q2$ corresponds to the jumping cue, calculated by the transition probability between "animal" and the next item in the fluency list. The Random Walk model is optimized by tuning the jump parameter ρ to weight the likelihood of stochastic "restart" required to produce the next item.

Fixed Count Model

The Fixed Count (FC) model is a patch-leaving decision rule from foraging theory which we adopt as a baseline model. The Fixed Count model dictates a switch after a fixed number of items have been captured in the current patch (Wilke, Hutchinson, Todd, & Czienskowski, 2009), and is best suited for environments in which there are constant numbers of items in each patch.

$$P(Switch|k) = \begin{cases} 1 & \text{if k = K} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$P(ContinueLocal|k) = 1 - P(Switch|k) \quad (5)$$

K corresponds to the fixed count value determining how many items are between each switch. The Fixed Count model is optimized by identifying the optimal fixed count value as a threshold to switch that best estimates each fluency list.

Equating Switch Models for Comparison

The Semantic Foraging model is based on a distributional semantic spatial representation, whereas the Random Walk model is based on a semantic network representation. As indicated in (Avery & Jones, 2018), it is an error to consider the two representations fundamentally different. Spatial and Networks representation are isomorphic in nature if they encode the same similarity data (e.g. same matrix representation). However, in the independent studies they were originally introduced, the Semantic Foraging and Random Walk models were based on different similarity data sources.

The original Semantic Foraging spatial model was built upon the cosine similarity matrix generated from the distributional semantic BEAGLE model (Jones & Mewhort, 2007).

The original Random Walk network model was based on the transition probability matrix estimated on a semantic network generated from word association data. To enable equal comparison between the two models, we perform analysis based on a similarity matrix generated from a pretrained Wikipedia Word2Vec model on items in the common dataset (Mikolov, Chen, Corrado, & Dean, 2013). Within the similarity matrix for the models, we also include the item "animal" to enable the jump cue of the Random Walk model. The Fixed Count Model does not depend on the underlying similarity data, and is directly equatable to both models for comparison.

Evaluating Model Performance

To evaluate the performance of each model for capturing switch predictions, we utilize the study data collected from the original Semantic Foraging paper by Hills et al (Hills et al., 2012). A total of 141 undergraduate students at Indiana University were recruited to participate in the study for course credit. Participants in the study were asked to participate in the semantic fluency task, entering as many animals as they could via keyboard, in 3 minutes. Each animal name and respective time of entry were recorded as items for downstream analysis. Each item entry was visible until "return" was entered by the participant, ensuring they could not see previous entries. Post-hoc hand-correction of spelling was applied and non-animal items were removed.

For each fluency list, the respective switch models were optimized to minimize the negative log likelihood of switching characterized by the extended Troyer Norms. The Troyer norms were used to classify transition points in the fluency lists via categorical shift in predetermined taxonomy of fluency items.

In the Semantic Scent model, λ was optimized per individual fluency list, while N was optimized globally across all transitions in the dataset. Though N can be optimized individually per fluency list, we opted for a global optimization scheme across all fluency lists for simplicity of implementation. Further individual optimization of N favors Semantic Scent performance in comparison to other switch models, as it would further minimize negative log likelihood of estimating each fluency list. This will be particularly important for characterizing data from clinical populations who produce more idiosyncratic switching behavior. In the Random Walk model, the jump parameter ρ was optimized per fluency list to estimate the jumping probability best characterizing the observed fluency list. In the Fixed Count model, K was optimized per individual list to best estimate the number of items between switches that best explains the fluency list.

Performance of each model was evaluated via application of the Bayesian Information Criterion (BIC) across each individual item transition observed in the dataset. In total, there were 5079 transitions in the dataset. The BICs are reported in **Table 1**. BIC is calculated according to:
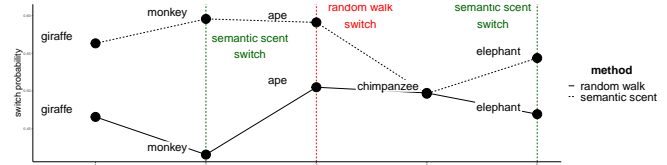
$$BIC = 2NLL + p \times log(N)$$

Figure 3: *Monkey*, *Ape*, and *Chimpanzee* consist of a retrieved item cluster. The Semantic Scent model predicts a switch at the beginning and after the cluster, while Random walk predicts a switch in the middle of the cluster at "Ape"

where *NLL* is the sum of the negative log likelihood across all transitions, p is the number of parameters optimized in each model, and N is the total number of transitions.

Qualitative Model Comparison

The key goal of each switching model is to predict accurate switching behavior in observed fluency data. We evaluate both the Semantic Scent and Random Walk models to compare their ability to identify cluster switches. Demonstrated in the original Semantic Foraging paper, switching behavior is best characterized by the reduction in residual proximity prior to switching, and spiking of residual proximity after switching (see Figure 3 from Hills et al. reproduced in **Figure 1**). Utilizing this heuristic, the Semantic Scent model appear to characterize norms-based switching behavior better the Random Walk model.

Hand-coded norms, including the Troyer norms, can be limited in their capacity for capturing clustering and switching behavior. However, switching models should be able to adequately predict switches based on stable category clusters in such hand-coded norms, given their predictive validity. We observed the performance of both Semantic Scent and Random Walk to perform switches in both examples seen in **Figures 3 and 4**. Switches are predicted when the probability of switching spikes, corresponding to the drop in residual proximity in the search space. In the example shown in Figure 3, *Monkey*, *Ape*, and *Chimpanzee* form a stable categorical cluster in the Troyer norms. Semantic Scent accurately predicted a switch between *Giraffe* and *Monkey*, corresponding to the start of the cluster, and predicts a switch between *Chimpanzee* and *Elephant*, at the end of the cluster. The Random Walk model also predicted a switch at *Ape* switch in the middle of a cluster.

In the example shown in Figure 4, *Ferret*, *Gopher*, and *Groundhog* form another stable categorical cluster in the Troyer norms. Semantic Scent predicted a switch between *Fish* and *Ferret*, at the start of the cluster. Random Walk predicted a switch between *Ferret* and *Gopher*, in the middle of a cluster. Random Walk further predicted a subsequent switch between *Gopher* and *Groundhog*, evidenced by an additional spike in switching probability.

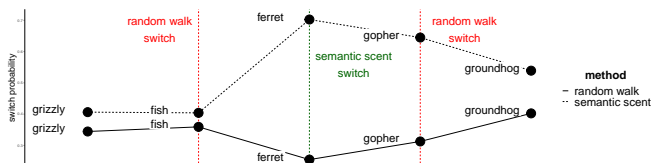In both examples, Semantic Scent correctly predicted the

Figure 4: *Ferret*, *Gopher*, and *Groundhog* consist of a retrieved item cluster. The Semantic Scent model predicts a switch at the beginning of the cluster, while Random Walk predicts a switch in the middle, and does not predict a switch between *Fish* and *Ferret*

start and end of a new norms-based cluster, whereas Random Walk predicted switches in the middle of clusters. Given the importance of predicting stable categorical clusters, the preliminary evidence suggests Semantic Scent may better characterize norms-based switching behavior than Random Walk.

Quantitative Model Comparison

In the examples of observed fluency data, Semantic Scent demonstrated better capture of switching behavior at the beginning and end of stable categorical clusters. Random Walk in such examples appeared to demonstrate more spurious switching behavior not aligned to the norms-based cluster categories. We further evaluate the performance of the models via calculating the Bayesian Information Criterion over the likelihood of all item transitions in the fluency dataset.

The Semantic Scent and Random Walk models strongly outperform the random comparison model, Fixed Count. This follows intuition, as cluster patches in the original dataset are not sized evenly, a key determinant for strong performance using the Fixed Count model. The Semantic Scent model also better minimizes the BIC compared to the Random Walk model, despite requiring two-fold parameter optimization which is penalized in calculating BIC . The BIC for all three models are reported in **Table 1**. We further calculated the Bayes Factor of the BIC difference between Semantic Scent and Random Walk. The resulting Bayes Factor ($K > 10^{12}$) suggests very strong evidence in favor of the Semantic Scent model, based on Jeffreys' interpretation of Bayes Factor (Jeffreys, 1935). This further validates our initial findings from example fluency data.

Table 1: Switching Model Results

Model	# of parameters	BIC
Semantic Scent	2	6903.93
Random Walk	1	7419.88
Fixed Count	1	34193.32

The strength of the Semantic Scent model in predicting switch behavior is likely due to its inspiration from scent-based models in animal ecology. The original findings in (Hills et al., 2012) suggested strong similarity between semantic memory search and animal foraging behav-

ior. As a result, it was implied that memory search mechanisms may have been exapted from the foraging mechanisms that evolved for searching for food resources in a physical space. The Semantic Scent mechanism simply extends the exploration-exploitation mechanism animal foraging behavior. Thus Semantic Scent is a natural extension of optimal foraging theory to memory search. Just as animals rely on the scent of food resource patches to determine when to switch to a new patch to optimize food foraging, Semantic Scent relies on the scent of item patches to switch between clusters in semantic memory to optimize semantic foraging.

Discussion

The present study sought to identify an optimal process model for predicting switching behavior in semantic memory search. Hills et al. (2012) identified a decrease in residual proximity as items were produced in a cluster, followed by a spike again when a switch occurred and a new semantically rich patch of items was located. The similarity-drop heuristic proposed in Hills et al. (2012) for classifying switch behavior provided a vital approach to characterize switches within the framework of the optimal foraging model. As influential as it was, the similarity-drop heuristic falls short as it is a postdictive rather than predictive indicator of switch. The Semantic Scent model was introduced in this paper to address this shortcoming, providing a predictive model of switching behavior in memory search. The Semantic Scent model directly addresses the observed relationship between residual proximity and cluster-and-switch behavior.

To evaluate it's effectiveness, we compared Semantic Scent to an analogous model, Random Walk. Random walk was shown to create the similar statistical signatures of foraging behavior with a simpler process model compared to optimal foraging, demonstrating effective parameterization of memory search behavior. In comparison to Random Walk, the Semantic Scent model appeared to provide a better explanation of switching behavior. The Semantic Scent model was able to predict cluster switches indicated by norms-based categories, whereas the Random Walk predicted some spurious switches within clusters. The behavior identified in these examples are further supported by quantitative evidence suggesting the strength of the Semantic scent model, demonstrated by the lower BIC and significant Bayes Factor value.

Given the superiority of the Semantic Scent model as a switching model to the comparison and baseline model, we will further interrogate its applicability to different applications of cluster and switch. To further validate our findings, we will compare the performance of Semantic Scent to comparison models on predicting participant-identified switches in their own fluency data. Further validation beyond predicting switches from the hand-coded Troyer norms will ensure the robustness of the switching model for potential downstream applications. A prime use case of the Semantic Scent model is evaluating of semantic memory in clinical populations. Given clustering and switching met-

rics are particularly sensitive to clinical group diagnoses (e.g. Alzheimer's, Parkinson's Disease, Schizophrenia), the Semantic Scent model can be utilized to estimate the likelihood of impairment or trajectory of disease in clinical populations.

There are also a few key assumptions of the Semantic Scent model that must be further assessed. First, the strong performance of the Semantic Scent model over the Random Walk model may be explained by the structural representation used. The current representation utilized in our work for both competing models is a distributional semantic space, estimated via cosine similarity between Word2Vec embeddings. The underlying semantic space is a generalized representation independent from experimental behavioral data. The original Random Walk model (Abbott et al., 2012) uses word association data to construct the underlying network. Word association data comes from experimental behavioral data, which may bias the results the Random Walk model on explaining SFT data (*e.g.* estimating behavior based on behavioral data). Thus, we seek to validate our methods on an isomorphic representation constructed from word association data, to determine if the strength of the Semantic Scent method in predicting switches generalizes to representations based on behavioral data. This would clarify whether the underlying representation affects the effectiveness of the competing models in characterizing SFT data.

Secondarily, the Semantic Scent model assumes an explicit switching mechanism in semantic memory search. The Random Walk model demonstrated that a simple process model was able to capture similar statistical signatures to a foraging-based model, without an explicit switching mechanism. However, the better performance of the Semantic Scent method in predicting norms-based switches suggests that an explicit switching mechanism may be important for characterizing semantic fluency. This necessitates further validation of switch predictions, such as comparing their ability to predict subject-determined switches (*e.g.* subject reports where they performed a switch) or switches based on the original similarity drop heuristic.

Overall, the present findings further solidify optimal foraging theory as an important explanatory process model for semantic memory search. After more rigorous evaluation of the assumptions of Semantic Scent are conducted, we seek to integrate the Semantic Scent switching model into the original Semantic Foraging model (Hills et al., 2012). Integration of Semantic Scent into Semantic Foraging allows us to produce a fully predictive model that can explain the cognitive mechanisms behind semantic memory search, moving the field beyond postdictive appraisal. This opens up opportunities to evaluate foraging methods on a larger-scale to better characterize memory search across all populations.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015, July). Random walks on semantic networks can resemble optimal foraging. *Psychological Review*, *122*(3), 558–569. doi: 10.1037/a0038693

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2012). Human memory search as a random walk in a semantic network. , 9.

Avery, J. E., & Jones, M. N. (2018). Comparing models of semantic fluency: Do humans forage optimally, or walk randomly? , 6.

Charnov, E. L. (1976, April). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. doi: 10.1016/0040-5809(76)90040-X

Chi, E. H., Pirolli, P., Chen, K., & Pitkow, J. (2001, March). Using information scent to model user information needs and actions and the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 490–497). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/365024.365325

Hills, T. T., Jones, M. N., & Todd, P. M. (2012, April). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. doi: 10.1037/a0027373

Hills, T. T., Todd, P. M., & Jones, M. N. (2015). Foraging in Semantic Fields: How We Search Through Memory. *Topics in Cognitive Science*, *7*(3), 513–534. doi: 10.1111/tops.12151

Jeffreys, H. (1935, April). Some Tests of Significance, Treated by the Theory of Probability. *Mathematical Proceedings of the Cambridge Philosophical Society*, *31*(2), 203–222. doi: 10.1017/S030500410001330X

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114*(1), 1–37. (Place: US Publisher: American Psychological Association) doi: 10.1037/0033-295X.114.1.1

Lundin, N. B., Todd, P. M., Jones, M. N., Avery, J. E., O'Donnell, B. F., & Hetrick, W. P. (2020, January). Semantic Search in Psychosis: Modeling Local Exploitation and Global Exploration. *Schizophrenia Bulletin Open*, *1*(1), sgaa011. doi: 10.1093/schizbullopen/sgaa011

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, September). Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781 [cs]*. (arXiv: 1301.3781)

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). *The PageRank Citation Ranking: Bringing Order to the Web.*

Pirolli, P., & Card, S. (1999). The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis. , 6.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93–134. (Place: US Publisher: American Psychological Association) doi: 10.1037/0033-295X.88.2.93

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, *4*(1), 28–34. (Place: United Kingdom Publisher: Blackwell Publishing) doi: 10.1111/j.1467-

9280.1993.tb00552.x

Shepard, R. N. (1987, September). Toward a Universal Law of Generalization for Psychological Science. *Science*, *237*(4820), 1317–1323. doi: 10.1126/science.3629243

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998, June). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, *36*(6), 499–504. doi: 10.1016/S0028-3932(97)00152-8

Wilke, A., Hutchinson, J. M. C., Todd, P. M., & Czienskowski, U. (2009). Fishing for the Right Words: Decision Rules for Human Foraging Behavior in Internal Search Tasks. *Cognitive Science*, *33*(3), 497–529. doi: 10.1111/j.1551-6709.2009.01020.x