An Agent-Based Model of Foraging in Semantic Memory

Diego Morales Bader
(dmoralesb@alumnos.uai.cl)
Faculty of Engineering and Science
Universidad Adolfo Ibáñez
Santiago, Chile

Enrique Canessa
(ecanessa@uai.cl)
Faculty of Engineering and Science
Universidad Adolfo Ibáñez
Viña del Mar, Chile

Sergio E. Chaigneau
(sergio.chaigneau@uai.cl)
Center for Social and Cognitive Neuroscience
School of Psychology
Universidad Adolfo Ibáñez
Santiago, Chile

Abstract

An agent-based model for semantic search and retrieval in memory is proposed. The model seeks to generate verbal fluency lists with properties similar to those generated by humans in the semantic fluency task. This model is compared to a random walk in a semantic network in its ability to adjust to the results of 141 undergraduate students in the semantic fluency task in eight different outcomes. We found that the agent-based model fits participants' results better than the random walk model. The results were consistent with optimal foraging theories, and the distributions of the total number of words, similarities, and frequency values were similar to those generated by participants. The potential uses of this model as a virtual environment to experiment with the search and retrieval process in semantic memory are discussed.

Keywords: Semantic Memory; Memory Retrieval; Semantic Fluency Task; Agent-Based Model; Optimal Foraging Theory

Introduction

It has been proposed that people search and retrieve items from their semantic memory following an optimal strategy of exploring and exploiting semantic space (i.e., not abandon a local patch too early or too late), similar to how nonhuman animals search for resources in physical space (Hills, Jones, & Todd, 2012; Hills, Todd, & Jones, 2015; Todd & Hills, 2020; Zemla, Gooding, & Austerweil, 2023). Humans may use optimal foraging strategies to search in memory, such as the marginal value theorem (MVT) (Charnov, 1976), which predicts a switch between patches when the marginal rate of return (the rate of finding new items per unit of time) in the current patch falls below the average rate of return for the entire search.

The MVT has been explored using the semantic fluency task (SFT), where participants generate words belonging to a semantic category in a limited time. Semantic fluency tasks are commonly used to assess executive functioning. The task has been successfully applied in studies on various neuropsychiatric disorders (Canning, Leach, Stuss, Ngo, & Black, 2004; Sebaldt et al., 2009). It is included in most neuropsychological assessment batteries (Sherman, Tan, & Hrabok, 2020) because it is considered a sensitive indicator of brain dysfunction.

Using the SFT, Hills et al. (2012) asked participants to retrieve as many animal names as possible in 3 minutes. They found that participants switched between patches of animal names in a way consistent with the MVT. Additionally, they found that participants with an optimal search process (according to the MVT) recalled more animal names than those who abandoned local search too quickly. They also explored the performance of different computational models, where the model that considered switching between local and global cues (cue-switching model) was the one that adjusted better to participants' fluency lists.

However, it has also been shown that results consistent with optimal foraging can be replicated with a simple Random Walk over a semantic network (Abbott, Austerweil, & Griffiths, 2015). Although the cue-switching model best fits participants' fluency lists, a Random Walk model also delivers the expected qualitative results when evaluated on the same semantic network representation (Avery & Jones, 2018).

These results suggest a discussion of the importance of search mechanisms relative to the importance of information structure and representation (Hills & Kenett, 2022). It can be argued that much of the process complexity is hidden in the data representation rather than in the intrinsic properties of search mechanisms (Jones, Hills, & Todd, 2015; Thompson & Kello, 2014). However, empirical data suggest that deviations from optimality (i.e., fewer words, fewer cluster changes, smaller cluster size) are related to impairments in cognitive mechanisms (e.g., executive functions) involved in search and retrieval, rather than impairments in representational structure (Bose, Wood, & Kiran, 2017; Tröger et al., 2019; Weakley & Schmitter-Edgecombe, 2014).

Despite considerable research on this issue, it is still unclear whether deficits in the SFT listing process reflect an impaired semantic memory or a preserved memory with an impairment in search and retrieval mechanisms (Duff, Covington, Hilverman, & Cohen, 2020; Rogers & Friedman, 2008; Weakley & Schmitter-Edgecombe, 2014).

To contribute to this debate, we propose an agent-based model that produces verbal fluency listings derived from the interaction between a search space and the control mechanisms related to the search process.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

A model with a good fit would allow us to explore hypotheses regarding how variations in parameters related to the regulation of the search process influence the model output for a given semantic structure. This study will focus on presenting the agent-based model and its capability to generate fluency lists with aggregate characteristics similar to lists produced by participants from a non-clinical sample.

Agent-Based Models

Agent-based models (ABMs) are computational models that simulate the behavior of individual agents and the interactions between them and their environment. One of its key characteristics is the bottom-up modeling approach, where emergent phenomena are produced through agents' interactions. Agents also can have adaptive behavior that allows them to change based on their experience.

ABMs are used in many fields, but not primarily in cognitive modeling of internal processes, except for some ACT-R models, where agents can be specified based on a defined cognitive framework (Anderson, Bothell, Lebiere, & Matessa, 1998; Pirolli & Card, 1999; Ritter, Tehranchi, & Oury, 2019). The reasons for this include historical, methodological, and practical discussions (Canessa, Chaigneau, & Moreno, 2023; Conte & Paolucci, 2014; Epstein, 1999). However, the benefits of broader adoption have also been highlighted because it would help us to better understand theories by modeling their components and interactions, provide us with a way to study a system that may be difficult or inaccessible to access, and for their value in being generative models of the variables of interest, among other reasons (Canessa et al., 2023).

Building an ABM of optimal foraging allows us to add more complexity than a Random Walk model to control for exploration and exploitation search in memory while keeping the phenomenon simple enough to evaluate the interactions between the most relevant mechanisms.

Using an ABM approach to generate data can be more advantageous than creating models that rely on existing data (e.g., the cue-switching model of Hills et al. (2012)). A generative approach allows for the construction of a virtual environment that can be used to study different mechanisms and their interactions, providing a way to analyze systems that may be difficult to get data from (such as specific clinical populations with cognitive impairments). Also, exploring this virtual environment could help devise new hypotheses for the memory search and retrieval process in different scenarios.

On the other hand, non-generative models that rely on existing data may not be able to account for interactions among different components of a system that may only be visible when producing the phenomenon. By conducting multiple experiments, the generative approach enables us to explore the effects of different parameters on search process outcomes. However, the first step is to ensure that the model generates the expected behavior, which is the main objective of this work.

The Model

The model consists of an agent that searches and retrieves words from a semantic space, similar to how animals forage in a physical space. Each agent has its own semantic space selected from a sample of a larger semantic space. The search process combines the cue-switching model of Hills et al. (2012) and the semantic scent model of Zhang and Jones (2022). In addition, mechanisms are incorporated to guide the search and encourage the exploration of semantic space.

In this model, clusters of words are not pre-defined; rather, they emerge from each agent's landscape structure and search process. This aspect relates to the variability between participants in the SFT. However, since the model returns a list of words, other methods to evaluate the clustering can be used, such as the hand-coded categorizations of animals from Troyer, Moscovitch, and Winocur (1997) or the similarity drop method from Hills et al. (2012), that identifies transitions when similarities decrease between words.

Model Landscape

Memory space is likely best represented in several dimensions; however, this model simplifies the landscape into a two-dimensional representation. The assumption is that for some word listing tasks like semantic fluency or word association tasks, a good approximation is possible in a two-dimensional space because higher-dimensional semantic information can be preserved in a lower dimensional representation (Croft, 2022; Lowe, 2001; Richie, White, Bhatia, & Hout, 2020).

One of the challenges of performing a two-dimensional representation is to capture the relationships between words that would form a cluster by proximity. For this purpose, we chose the t-distributed stochastic neighbor embedding (t-SNE). This statistical non-linear method tries to find a way to project data into a low-dimensional space so that the clustering in the high-dimensional space is preserved (Van der Maaten & Hinton, 2008). t-SNE maintains the neighborhood data points closer, preserving the local structure more than the global structure of the data.

Since t-SNE is a non-deterministic (randomized) algorithm, the landscape presents some variability between agents even if the same words are selected to populate the landscape. This variability is also probably present in the case of empirical data.

To build the landscape, we used the semantic similarity matrix (cosine similarity) and frequency of animal names ($n = 765$) from Hills et al. (2012), which was constructed from a Wikipedia corpus using the learning model BEAGLE (Jones & Mewhort, 2007). For each agent, a sample of words is probabilistically selected, where words with higher frequency are more likely to be selected. An angular distance matrix is calculated from the cosine similarity matrix of the selected words, by which the t-SNE algorithm generates the two-dimensional representation of the data, as seen in Figure 1.

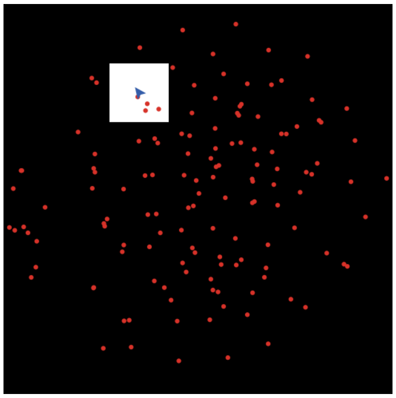

Figure 1: Example of a landscape with 138 words constructed with t-SNE. The dots are the words that can be retrieved. The blue triangle is the agent. The white square is the agent's neighborhood where it can detect words.

A word is selected with a probabilistic selection based on logarithm frequency values to decide the agent's initial position in the landscape. The agent is placed near the selected word.

Search Process

The search process is based on Hills et al. (2012) cue-switching model, which considers a local search that uses the similarity and frequency values of the words as cues and a global search that considers only the frequency of the words. Their model requires a list of words already produced to know which cues to use and does not indicate when the transition between search types should occur. Also, they fit a saliency parameter to adjust the likelihood of the model to the participant's lists.

We removed the saliency parameter for our model because the generative process does not rely on already produced lists. In addition, instead of considering the entire lexicon, the evaluation is performed on words found in the agent's immediate neighborhood (see white square in Figure 1).

When the agents are in local search, a probabilistic selection is performed between the words in the agent's neighborhood, where the weights are the cosine similarity values related to the last word reported by the agent multiplied by the logarithm of the candidate word frequency value. When the agents perform a global search, the selection is based only on the logarithm of frequency values as weights (Hills et al., 2012).

We employ an equation based on the "semantic scent" model proposed by Zhang and Jones (2022) to define when the agent must transition between search types. The likelihood of switching from a patch based on the item just produced, and the proximity to neighborhood items is given by

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

$$P(\text{ContinueLocal}|X,N) = 1 - P(\text{Switch}) \tag{2}$$

where Y corresponds to the vectors of N most similar items to X in the current agent neighborhood. Y_i corresponds to the i^{th} most similar item to the current item X. λ is a sensitivity parameter to control the shape of the function.

Equation 1 adapts the equation of the "semantic scent" in Zhang and Jones (2022) so that when $\lambda = 0$, the probability of switching is also zero (agents always deplete their current patch). When λ increases, the probability of switching approaches to 1 (agents tend to leave their current patch too early).

In our model, Equation 1 defines the probability at each moment of switching from local search to global search based on similarity values and the probability of switching from global search to local search based on the normalized frequency values (between 0 and 1). Typically, the probability of leaving the local patch increases when the patch value and the number of words decrease, and the probability of entering a patch is higher when there are many words or the words have high-frequency values. A probabilistic selection determines if the agent should switch between search types.

When the agent reports a word, that word is not considered again to compute the switching probability and is not reported again. This procedure is a simplification of a control mechanism for checking errors and avoiding repetitions. For example, the model Search of Associative Memory (SAM) assumes that an item previously recalled cannot be recalled again within a given trial (Raaijmakers & Shiffrin, 1981, 1980).

Also, when the agent leaves a patch or fails to enter a patch, if there are remaining words in the neighborhood, those words are not considered again in any probability computation or selection unless some of those words are in the agent neighborhood after the decision to enter a patch and start a local search is made.

Agent movement

Agent movements are a combination of goal-directed movement and random walk. When a word is selected to be reported, the agent moves to it in a straight line, one simulation step (tick) at a time, until it reaches the word. When the agent is in global search, the movement is a random walk; however, if the probability in Equation 1 is greater than at the agent's last position, the agent continues to move in that direction. In other words, if the richness of the agent neighborhood in the current step is greater than the richness in the previous step, the agent continues to move in the same direction as in the previous step instead of performing the random walk. Thus, the agent moves in a positive gradient of word richness.

In global search, the neighborhood moves with the agent in each step so that the agent is always at the center. Once a decision to report a word is made, the neighborhood remains in the same position while the agent moves inside it. The neighborhood moves with the agent again after exiting the local search.

The agent can store and avoid empty patches if a memory parameter is set. If only empty patches are already in memory as movement options, the agent performs a random walk until a new patch is found. A limit can be set to store empty patches. If the limit is exceeded, the oldest empty patch in memory is deleted to make space for the new empty patch.

Parameter Fitting

We compared the lists generated by our model with those produced by the participants from Hills et al. (2012) (141 undergraduates), with the same similarity matrix and frequency values from the animal category. The same set of analyses done by Hills et al. (2012) were performed: a) the similarity between a word and the words preceding it, b) the ratio of pairwise similarity over the subject's mean similarity by patch entry position, c) the residual proximity (mean similarity to all possible remaining words) of an item to an item's position before or after a patch transition, and d) the mean ratio between the inter-item retrieval time (IRT) for an item and the participant's mean IRT over the entire task, relative to the order of entry for the item.

Additionally, we compared the distribution of numbers of words, similarity and frequency values produced, the average number of patches, and the average patch size. In summary, we have eight different variables to evaluate the model outcomes.

Our model was compared with the Random Walk over a semantic network proposed by Abbott et al. (2015), which generates results consistent with the MVT. We set up the semantic network with weights based on the threshold method used by Avery and Jones (2018). This threshold indicates which similarity values will form a link between nodes. After this, a node "animal" is created, which connects to every node using the frequency value of each word as a link. For both models, 141 fluency lists were produced. The IRT computation was made following the Avery and Jones (2018) method, which computes the time (steps or ticks) to produce a non-repeated word and adds the number of letters of that word to that time (to reflect that words were typed on a keyboard when the original data were collected).

To determine when a patch transition occurred in a way that can be compared to the data of Hills et al. (2012), we adjusted the parameters with two methods: the hand-coded categorization norms of animals from Troyer et al. (1997) and extended by Hills et al. (2012), and the similarity drop method from Hills et al. (2012), that identifies transitions when similarities drop between words and rise again in the following word.

For the parameter optimization of the models, we used the Simulated Annealing method implemented in NetLogo's BehaviorSearch (Stonedah, 2010). The function to minimize was the sum of each variable's normalized root mean square error (NRMSE) (normalized by the range of the comparison values). A histogram with ten bins was computed to calculate the NRMSE for the number of words, similarity, and frequency distributions. The average number of patches and patch size were considered together while calculating the NRMSE. The only fixed parameters were the patch memory size and the neighborhood radius for the semantic foraging ABM.

In addition, we wanted to penalize models for the number of parameters; however, determining the complexity of ABMs and penalizing them is non-trivial (Lux & Zwinkels, 2018; Mandes & Winker, 2017). Methods such as the Akaike Information Criterion (AIC) (Akaike, 1974) and Bayesian Information Criterion (BIC) (Schwarz, 1978) cannot be directly applied in our case since we do not have a closed-form probabilistic equation describing our model where we can apply the maximum likelihood estimation method for our parameters. However, we approximate the AIC and BIC by treating the model results on all variables and participant data as a linear regression problem. For this, we normalized the model and participant results, and used the mean square error instead of the log-likelihood to calculate an approximation to the AIC and BIC. We used the AIC correction for small sample sizes (AICc) (Sugiura, 1978). We stress that this is only an approximate and preliminary method for comparing and selecting ABMs.

Both models were implemented in NetLogo 6.4.0 (Wilensky, 1999) and can be downloaded from `https://osf.io/qsz2e/` along with the data analysis script (Python), the pseudocode and a flowchart of the model, and tables with the descriptions of the parameters and the best fitting values found by the Simulated Annealing method.

Results

Using the best-fit parameters to the Hills et al. (2012) data, 50 runs were performed (141 lists of words for each run) for each model and each patch transition method (Troyer norms and similarity drop). It was found that independent of the patch transition method, the semantic foraging ABM has a lower sum of NRMSE ($M = 2.94$, $SD = 0.12$) than the random walk over a semantic network ($M = 5.68$, $SD = 0.26$), $t(198) = -16.50$, $p < .0001, \eta^2 = 0.58$. In addition, both models obtained a better fit with the similarity drop method ($M = 3.24$, $SD = 0.10$) of Hills et al. Hills et al. (2012) compared to Troyer's norms ($M = 5.37$, $SD = 0.28$), $t(198) = -10.33$, $p < .0001$, $\eta^2 = 0.35$. Figure 2 shows the pairwise comparisons between models and within the patch transition methods.

When penalizing for the number of parameters, the foraging ABM, which has five additional parameters, obtained on average a lower approximate AICc ($-178.18 < -131.78$) and lower approximate BIC ($-167.78 < -121.07$) compared to the random walk model, both with the similarity drop method. With Troyer's norms, our model (AICc $= -150.26$, BIC $= -139.54$) also obtained a lower value than the random walk model (AICc $= -115.59$, BIC $= -110.37$). While this would indicate that the foraging ABM should be chosen, this method of calculating AICc and BIC is only a rough approximation, so these results should be taken cautiously.

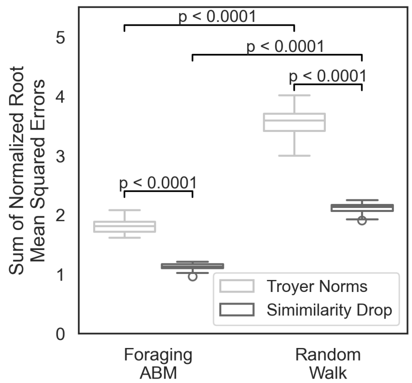

Figure 2: Sum of the NRMSE of the semantic foraging ABM and the random walk over a semantic network for each patch transition method. For each model, 50 runs were performed with the best parameters found for each patch method.

The fit of the distributions generated by the models to the participant's data was evaluated using the chi-square distance to the participant's distribution of the number of words, similarity, and the logarithm of the frequency of words. Ten equivalent bins were used between the participants' data and those generated by the models. It was found that, on average, considering 50 runs of both models (with 141 fluency lists in each run), the semantic foraging ABM has a significantly smaller distance to participants' distributions in the number of words produced ($t(98) = -28.77$, $p < .0001, d = 5.75$), similarity ($t(98) = -100.42$, $p < .0001, d = 20.08$), and frequency ($t(98) = -120.30$, $p < .0001, d = 24.06$) compared to the random walk model. Figure 3 shows an example of the distributions of the models and the participant's distributions. It can be qualitatively observed that the semantic foraging model produces distributions closer to those produced by the participants.

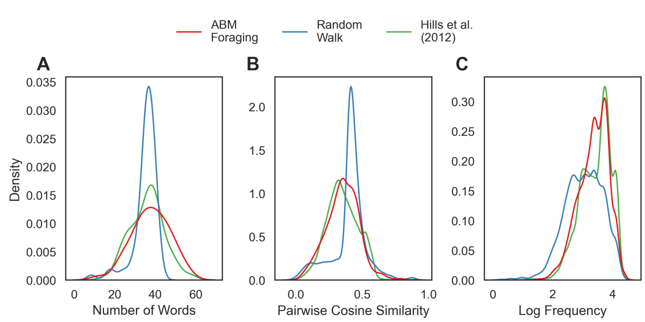

Figure 3: KDE plot of the number of words (panel A), pairwise cosine similarity (panel B), and frequency logarithm (panel C) of the words produced in one run with the best parameters.

Additionally, the marginal value theorem prediction that consistent patch departure times result in more recovered words was evaluated. One sign of consistency is the last IRT in a patch being close to the mean global IRT. If that last IRT is shorter, it is a sign of leaving patches too soon, and if it is much longer, it is a sign of staying too long.

The similarity drop method was used to compute the difference between mean last-item IRT across patches and overall task IRT. Linear regression was performed for each model to predict the number of items produced (Figure 4). Normalized differences were analyzed after removing outliers more than three standard deviations from their respective means. A significant negative relationship (longer last IRT deviations from mean IRT led to fewer words retrieved) was found for the semantic foraging ABM, $\beta = -25$, $t(136) = -8.05$, $p < .0001 (R^2 = 0.32)$, and the random walk model, $\beta = -7.2$, $t(137) = -3.36$, $p = .0001 (R^2 = 0.08)$, however, the tendency was more robust in the semantic foraging ABM.

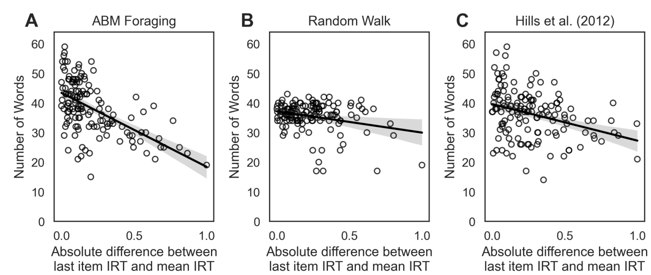

Figure 4: The relationship between deviation from the marginal value theorem policy for patch departures with similarity drop method (horizontal axis) and the total number of words produced in each model and participants' data. The line is the best-fitting linear regression.

Finally, the prediction that the optimal time for a patch switch should happen when the current intake rate in the patch falls to the mean global intake rate for all patches was evaluated. This statement is evaluated using the inter-item retrieval time (IRT) at patch switches relative to the mean IRT across all words produced. The IRTs are considered as the intake rate.

Using the similarity drop method, we found that the word immediately following a patch switch takes significantly longer to produce on average than the mean IRT for both models, $t(1492) = 7.22$, $p < .0001$, $d = 0.19$ (semantic foraging ABM), $t(1542) = 6.66$, $p < .0001$, $d = 0.17$ (random walk). The second word in a patch takes significantly less time than the mean IRT for both models, $t(1457) = -7.74$, $p < .0001$, $d = 0.20$ (semantic foraging ABM), $t(1486) = -4.49$, $p < .0001$, $d = 0.12$ (random walk). These results show that patch transitions IRTs align with the marginal value theorem for both models, as seen in Figure 5.

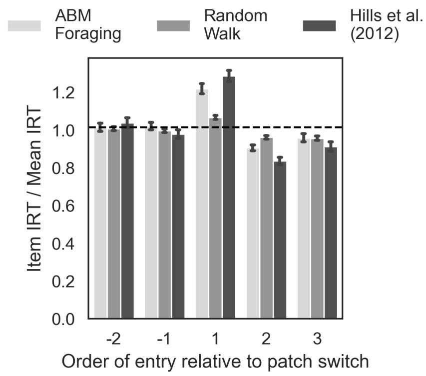

Figure 5: The mean ratio between the inter-item retrieval time (IRT) and the mean IRT relative to the order of entry of the patch (similarity drop method). The bars above "1" indicate the relative IRT between the first word in a patch and the last word in the preceding patch. The dotted line is the participant's mean IRT. Error bars are standard errors of the mean.

Discussion

We tested an agent-based semantic search model by evaluating its ability to generate word lists whose aggregated patterns are similar to those produced by people performing the verbal fluency task in eight different outcomes. The model used an area-restricted search, transitioning between local patch exploitation and global search. We compared our model with a random walk over a semantic network. Word information (similarity and frequency) was obtained from the BEAGLE learning model used by Hills et al. (2012).

While both models delivered results consistent with the marginal value theorem, we found that our model obtained a better fit than the random walk on a semantic network. These results are in line with the findings of Avery and Jones (2018), who compared the cue-switching model of Hills et al. (2012) with the random walk model of Abbott et al. (2015) and found that the cue-switching model fits the participants' fluency lists better. Since our model was based on the cue-switching model, we also expected a better adjustment. Moreover, in our model, the distributions of the number of words, similarity, and frequency were closer to those obtained by the participants compared to the lists generated by the random walk model.

Additionally, we tested two patch transition identification methods, Troyer et al. (1997) hand-coded norms and Hills et al. (2012) similarity drop. We found a better adjustment for both models with the similarity drop method. This result is expected since both models rely on the raw cosine similarity and frequency information obtained from a learning model without any adjustment for words more likely to be produced by participants. In other words, since the information obtained from the learning model (BEAGLE) and the information derived from human-produced lists are significantly different, the worst fit of the cluster transitions with the Troyer norms was expected (Johns & Jones, 2010).

We acknowledge that our model has more parameters than the random walk model, making it more likely to exhibit a better fit. However, we calculated approximate indices (AICc and BIC) that penalize our model's complexity and found that our model favorably compared to the random walk model. We reiterate that using AICc and BIC in this context is debatable. Evaluating ABMs complexity is a non-trivial task (Lux & Zwinkels, 2018; Mandes & Winker, 2017), and qualitative and theoretical criteria should also be considered. We argue that our model parameters have theoretical value in a virtual environment that seeks to simulate different word listing scenarios and offers more alternatives in the range of theory-based behaviors that can be experimented with in this environment.

In summary, we showed that our ABM has the potential to generate word lists that resemble human data patterns. However, further analysis is required to determine the degree of influence of each parameter on the results of interest (sensitivity analysis) and to evaluate the model generalization to other datasets. With a good generalization, the model can be a useful tool to test hypotheses on the interplay between memory structure and control mechanisms of the search process. For example, the next steps in the development of the model are to explore whether forcing agents to leave their patches earlier or constantly depleting their local patches can resemble results from populations with cognitive impairments (Arán Filippetti, Krumm, & López, 2023; Bose et al., 2017; Tröger et al., 2019). In contrast, we will explore if the structure of the semantic space is the main contributor to explaining optimal memory retrieval by varying the distribution of the distances between words. Finally, we will implement a Lévy flight search process (Rhodes & Turvey, 2007) in the same two-dimensional semantic space implemented in our model to make model comparisons in the same spatial representation. Furthermore, this model could be flexible enough to generate data of any semantic category type. We plan to extend the model to the Property Listing Task, where some mathematical properties have been found to differ from the SFT (Canessa & Chaigneau, 2020).

One limitation of our model comparison approach is that the differences between the two models might not come from the search process but from the decision to include variability in the semantic space. Further analysis and a new implementation of the random walk model with semantic network variability are required to determine what variables are producing the main differences found in our results.

Another limitation is that our model cannot adjust individual subjects' parameters. The model is designed to generate fluency lists with similar characteristics but not identical to the participant's produced lists. The lists generated by the model should be considered as if a different group of participants produced them. To fit the model for each subject, we must estimate the most probable semantic space and obtain the most probable trajectory for each subject. Methods to perform this fitting will be explored in the future.

Acknowledgments

This work received financial support from a James S. McDonnell Foundation grant *"A method for studying shared meaning in cultural groups"* `https://doi.org/10.37717/2022-3930`. Diego Morales Bader is supported by a doctoral scholarship from the Agencia Nacional de Investigación y Desarrollo (ANID) of the Government of Chile, Grant #21220612.

References

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

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. doi: 10.1109/TAC.1974.1100705

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An Integrated Theory of List Memory. *Journal of Memory and Language*, *38*(4), 341–380. doi: 10.1006/jmla.1997.2553

Arán Filippetti, V., Krumm, G., & López, M. B. (2023). Clustering and Switching During Verbal Fluency in Typical and Atypical Development: A Systematic Review in Children and Adolescents. *Current Psychology*, *42*(31), 27516–27533. doi: 10.1007/s12144-022-03787-4

Avery, J. E., & Jones, M. N. (2018). Comparing models of semantic fluency: Do humans forage optimally, or walk randomly? In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society* (pp. 118–123). Austin, TX: Cognitive Science Society.

Bose, A., Wood, R., & Kiran, S. (2017). Semantic fluency in aphasia: Clustering and switching in the course of 1 minute: Semantic fluency in aphasia. *International Journal of Language & Communication Disorders*, *52*(3), 334–345. doi: 10.1111/1460-6984.12276

Canessa, E., & Chaigneau, S. E. (2020). Mathematical regularities of data from the property listing task. *Journal of Mathematical Psychology*, *97*, 102376. doi: 10.1016/j.jmp.2020.102376

Canessa, E., Chaigneau, S. E., & Moreno, S. (2023). Describing and understanding the time course of the property listing task. *Cognitive Processing*. doi: 10.1007/s10339-023-01160-2

Canning, S. D., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology*, *62*(4), 556–562. doi: 10.1212/WNL.62.4.556

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

Conte, R., & Paolucci, M. (2014). On agent-based modeling and computational social science. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.00668

Croft, W. (2022). On two mathematical representations for "semantic maps". *Zeitschrift für Sprachwissenschaft*, *41*(1), 67–87. doi: 10.1515/zfs-2021-2040

Duff, M. C., Covington, N. V., Hilverman, C., & Cohen, N. J. (2020). Semantic Memory and the Hippocampus: Revisiting, Reaffirming, and Extending the Reach of Their Critical Relationship. *Frontiers in Human Neuroscience*, *13*, 471. doi: 10.3389/fnhum.2019.00471

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, *4*(5), 41–60. doi: 10.1002/(SICI)1099-0526(199905/06)4:5⟨41::AID-CPLX9⟩3.0.CO;2-F

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

Hills, T. T., & Kenett, Y. N. (2022). Is the Mind a Network? Maps, Vehicles, and Skyhooks in Cognitive Network Science. *Topics in Cognitive Science*, *14*(1), 189–208. doi: 10.1111/tops.12570

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

Johns, B. T., & Jones, M. N. (2010). Evaluating the random representation assumption of lexical semantics in cognitive models. *Psychonomic Bulletin & Review*, *17*(5), 662–672. doi: 10.3758/PBR.17.5.662

Jones, M. N., Hills, T. T., & Todd, P. M. (2015). Hidden processes in structural representations: A reply to Abbott, Austerweil, and Griffiths (2015). *Psychological Review*, *122*(3), 570–574. doi: 10.1037/a0039248

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. doi: 10.1037/0033-295X.114.1.1

Lowe, W. (2001). What is the Dimensionality of Human Semantic Space? In R. M. French & J. P. Sougné (Eds.), *Connectionist Models of Learning, Development and Evolution* (pp. 303–311). London: Springer London. doi: 10.1007/978-1-4471-0281-6_30

Lux, T., & Zwinkels, R. C. (2018). Empirical Validation of Agent-Based Models. In *Handbook of Computational Economics* (Vol. 4, pp. 437–488). Elsevier. doi: 10.1016/bs.hescom.2018.02.003

Mandes, A., & Winker, P. (2017). Complexity and model comparison in agent based modeling of financial markets. *Journal of Economic Interaction and Coordination*, *12*(3), 469–506. doi: 10.1007/s11403-016-0173-0

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643–675. doi: 10.1037/0033-295X.106.4.643

Raaijmakers, J. G., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search in associative memory. *Psychology of Learning and Motivation*, *14*, 207–262. doi: 10.1016/s0079-7421(08)60162-0

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

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A: Statistical Mechanics and its Applications*, *385*(1), 255–260. doi: 10.1016/j.physa.2007.07.001

Richie, R., White, B., Bhatia, S., & Hout, M. C. (2020). The spatial arrangement method of measuring similarity can capture high-dimensional semantic structures. *Behavior Research Methods*, *52*(5), 1906–1928. doi: 10.3758/s13428-020-01362-y

Ritter, F. E., Tehranchi, F., & Oury, J. D. (2019). ACT-R: A cognitive architecture for modeling cognition. *WIREs Cognitive Science*, *10*(3), e1488. doi: 10.1002/wcs.1488

Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia*, *46*(1), 12–21. doi: 10.1016/j.neuropsychologia.2007.08.010

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2). doi: 10.1214/aos/1176344136

Sebaldt, R., Dalziel, W., Massoud, F., Tanguay, A., Ward, R., Thabane, L., … Lescrauwaet, B. (2009). Detection of Cognitive Impairment and Dementia Using the Animal Fluency Test: The DECIDE Study. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*, *36*(5), 599–604. doi: 10.1017/S0317167100008106

Sherman, E. M. S., Tan, J. E., & Hrabok, M. (2020). *A compendium of neuropsychological tests: Administration, tests, norms, and commentary* (Fourth edition ed.). New York: Oxford University Press.

Stonedah, F. (2010). *BehaviorSearch (1.10) [Computer software]*. http://www.behaviorsearch.org/.

Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections: Further analysis of the data by akaike' s. *Communications in Statistics - Theory and Methods*, *7*(1), 13–26. doi: 10.1080/03610927808827599

Thompson, G. W., & Kello, C. T. (2014). Walking across Wikipedia: A scale-free network model of semantic memory retrieval. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.00086

Todd, P. M., & Hills, T. T. (2020). Foraging in Mind. *Current Directions in Psychological Science*, *29*(3), 309–315. doi: 10.1177/0963721420915861

Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J., Peter, J., & Kray, J. (2019). Exploitation vs. exploration—computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. *Neuropsychologia*, *131*, 53–61. doi: 10.1016/j.neuropsychologia.2019.05.007

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138–146. doi: 10.1037/0894-4105.11.1.138

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(11), 2579–2605.

Weakley, A., & Schmitter-Edgecombe, M. (2014). Analysis of Verbal Fluency Ability in Alzheimer's Disease: The Role of Clustering, Switching and Semantic Proximities. *Archives of Clinical Neuropsychology*, *29*(3), 256–268. doi: 10.1093/arclin/acu010

Wilensky, U. (1999). *Netlogo.* Center for Connected Learning and Computer-Based Modeling. Northwestern University.

Zemla, J. C., Gooding, D. C., & Austerweil, J. L. (2023). Evidence for optimal semantic search throughout adulthood. *Scientific Reports*, *13*(1), 22528. doi: 10.1038/s41598-023-49858-9

Zhang, L., & Jones, M. N. (2022). Using "Semantic Scent" to Predict Item-Specific Clustering and Switching Patterns in Memory Search. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (pp. 3236–3242).