UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Mouse-mole-vole: The inconspicuous benefit of phonology during retrieval from semantic memory

Permalink

https://escholarship.org/uc/item/4164178j

Authors

Kumar, Abhilasha A
Lundin, Nancy B
Jones, Michael N.

Publication Date

2022

Peer reviewed

Mouse-mole-vole: The inconspicuous benefit of phonology during retrieval from semantic memory

Abhilasha A. Kumar
Indiana University
kumaraa@iu.edu

Nancy B. Lundin
Indiana University
The Ohio State University
nancy.lundin@osumc.edu

Michael N. Jones
Indiana University
jonesmn@iu.edu

Abstract

When concepts are retrieved from memory, this process occurs within a rich search space where multiple sources of information interact with each other. Although the mapping from wordform to meaning is generally considered to be arbitrary, there is recent evidence to suggest that form and meaning may be correlated in natural language, and semantic and phonological cues may interact during retrieval. However, whether phonology interacts with meaning-related information in deeper semantic retrieval tasks, and whether this interaction has broader implications for how we conceptualize semantic retrieval remains relatively understudied. We examined these questions within the framework of the semantic fluency task, where individuals were asked to retrieve as many exemplars as they could from a given category (e.g., animals) within a fixed period of time. Responses were more phonologically similar during later stages of retrieval, and greater phonological similarity across responses was associated with greater number of items produced. We formulated a nested set of optimal foraging models to evaluate the combined influence of semantic and phonological information on retrieval likelihood. Model comparisons revealed that a model that combined frequency, semantic, and phonological information locally to make within-category transitions but relied on only frequency as a global cue to make between-category transitions produced the best explanation of the behavioral data.

Keywords: semantic retrieval; verbal fluency; memory search; distributional semantics; cognitive model

Introduction

Retrieval of concepts from memory is ubiquitous - we retrieve concepts when we make grocery lists, pack items for a vacation, recall our favorite movies, and so on. Much of this retrieval is mediated via "semantic memory" - our storehouse of information about the meaning of objects, facts, and events. Indeed, semantic memory retrieval continues to be widely studied through theoretical (Kumar et al., 2021; Lake & Murphy, 2021) and computational perspectives (De Deyne et al., 2019; Hills et al., 2012; Davelaar, 2015), due to its far-reaching implications for broader cognition.

In many of these accounts, meaning is studied as an independent system, and the mapping between meaning and wordform is generally considered to be arbitrary. However, there is some evidence to suggest that form and meaning may indeed be correlated in natural language, due to functional pressures associated with language use and acquisition (Dingemanse et al., 2015; Monaghan et al., 2011). For example, it may be easier to learn the relationship between two phonologically related words in Hindi, *daayein* and *baayein* if

they refer to semantically related entities *right* and *left*, compared to two semantically unrelated words such as *ball* and *telephone*. Indeed, confirming this intuition, in a large cross-linguistic analysis of 100 languages, Dautriche et al. (2017) found a modest but reliable correlation between the phonological forms of words and their meanings (also see Blasi et al., 2016). Besides the functional advantage of this non-arbitrary mapping between form and meaning in natural language, a relatively unexplored question is whether semantic retrieval is sensitive to this mapping, and if so, does accessing the form-meaning relationship facilitate or inhibit retrieval in deeper semantic tasks?

A widely used semantic retrieval task is the semantic fluency task (SFT), where individuals are asked to produce as many exemplars as possible from a given semantic category (e.g., animals) within a fixed period of time. A large body of work has examined the search and retrieval processes involved in SFT, due to its broader implications for healthy (Hills et al., 2012, 2015) and impaired cognition (Lundin et al., 2020; Troyer et al., 1998). Responses in the fluency task are typically clustered into subordinate categories (e.g., pets, farm animals, etc.) and although participants sometimes produce phonologically related items (e.g., cat-bat, mouse-monkey, etc.), the interplay of semantic and phonological information in the fluency task has not been systematically studied.

Hills et al. (2012) tested a variety of search models applied to a semantic space simulated by a corpus-based distributional model of semantic representation, BEAGLE (Jones & Mewhort, 2007). The specific search model that best explained the human data was a dynamic two-cue model that used local similarity to generate items until no other proximal item was found, and then switched to a global frequency cue to select the next item (after which search by local similarity resumed). The fact that the global-local switch model produced the best fit to the human data was theoretically significant because it mirrored the best accounts of how animals make exploration-exploitation decisions when foraging for food. Just as a honey bee 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 "patch" of farm animals and search for a new resource-rich semantic patch (e.g., pets) to exploit. However, the search

space in these models considered only semantic similarity and frequency, not phonological cues; phonology was presumed irrelevant to a semantic retrieval task. The Hills et al. work was well before Dautriche et al. (2017) had demonstrated wordform-meaning correlations in language on a large scale, and the role of phonology in SFT needs to be fully explored.

One possibility is that due to phonological information being an irrelevant source of information when producing semantically related exemplars, it would have no significant influence on retrieval performance in the fluency task. Another possibility is that phonological information, due to being implicitly correlated with semantic information in natural language, is in fact used as an additional retrieval cue. If so, it remains unclear whether this additional cue, "irrelevant" to semantic search, is useful or causes interference during retrieval. On one hand, phonology may be a *global* cue, which may be useful in guiding the individual to a new semantic cluster when the current cluster is sufficiently depleted. On the other hand, phonology may be a *local* cue like semantic similarity that is used to identify items that are close within a given neighborhood. Finally, it is possible that phonological information acts as a "lure" and leads to poorer retrieval performance overall.

In this paper, we attempted to discriminate between these possibilities by investigating whether responses in the semantic fluency task show any evidence for phonological clustering, and whether this has implications for retrieval performance. Specifically, we analyzed two datasets of semantic fluency and statistically compared a series of optimal foraging models with different weights for semantic and phonological information, to evaluate the relative contribution of these cues to semantic retrieval [1].

Method

Data

We used two datasets of semantic fluency. The first dataset (hereafter referred to as the HJT dataset) was collected by Hills et al. (2012) from 141 participants, who were asked to retrieve as many exemplars as they could within 3 minutes for different semantic categories. The task was computerized and participants typed in their responses.

We also analyzed a more recent dataset of 30 participants from Lundin (in press) and Lundin et al. (in preparation), hereafter referred to as the LEA dataset. In this study, participants verbally produced fluency responses for 3 minutes for different semantic categories while undergoing fMRI scanning. Critically, after generating items in the fluency task, participants were shown a spreadsheet of their responses by the experimenter and asked to group items based on their subjective estimate of relatedness across items. These groupings were then coded as *participant-designated switches*. We utilized this coding scheme to better understand how individuals

clustered their responses as a function of semantic and phonological similarity (see Results section). We analyzed these participant-designated switches because Lundin (in press) recently demonstrated that these switches were sensitive to differences in functional brain activity when participants were switching compared to clustering.

Unless noted otherwise, analyses are reported on the concatenated data from the HJT and LEA datasets given that participants in both studies performed the same task (SFT). Furthermore, we only analyzed the *animals* category in this paper, given that the majority of work in this domain has focused on the *animals* category, and this category also has the most extensive hand-coded norms (Troyer, 2000).

Response Variables

Semantic Similarity We computed the semantic similarity between consecutive responses produced by each participant as the cosine similarity in a distributional semantic space, word2vec (Mikolov et al., 2013), trained on the Google news corpus, available via Patel et al. (2018). Figure 1 displays the consecutive semantic similarity between the responses for the response sequence giraffe→buffalo→bison→lion→tiger.
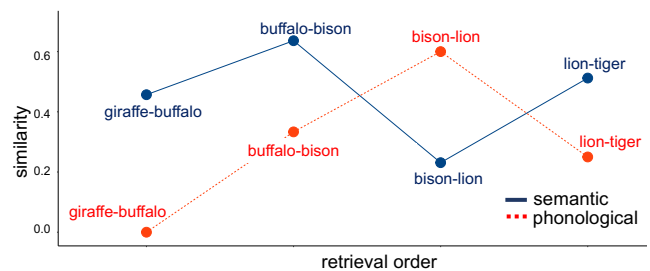

Figure 1: Semantic and phonological similarity across successive responses produced during the fluency task by a participant: semantic similarity is high between buffalo→bison and lion→tiger, but phonological similarity is high between bison→lion

Phonological Similarity We computed the phonological similarity between consecutive responses as the normalized edit distance between the phonemes contained within the responses. Phonemes were obtained using the CMU Pronouncing Dictionary, which maps words to their pronunciations based on the ARPAbet phonetic transcription. Normalized edit distance was computed as follows:

$$d(a,b) = 1 - \frac{e(a,b)}{\max(l_a, l_b)} \quad (1)$$

where e(a,b) signifies the edit distance between two strings of phonemes, *a* and *b*, and l_a and l_b denote the lengths of *a* and *b*. Figure 1 displays the consecutive phonological similarity between the responses for the response sequence giraffe→buffalo→bison→lion→tiger.

Word Frequency For each response, we also obtained the log-frequency derived from the SUBTLEX corpus, available via the English Lexicon Project (Balota et al., 2007).

Behavioral Results

Response statistics

Participants produced an average of 36.8 (*SD*=8.52) animal responses in the HJT dataset and 48.3 (*SD*=11.9) responses in the LEA dataset. Table 1 displays the mean semantic and phonological similarity of responses produced across both datasets, averaged across all participants.

dataset	semantic similarity	phonological similarity
HJT	.43 (.16)	.07 (.12)
LEA	.43 (.16)	.09 (.13)

Table 1: Mean semantic and phonological similarity in the fluency datasets

Cue usage over time

First, we examined whether the usage of semantic and phonological cues varied over the duration of the task. Figure 2 displays the mean semantic and phonological similarity as a function of the retrieval order, collapsed across both datasets. As shown, semantic similarity declined over the course of the task, whereas phonological similarity increased during the course of the task. Consistent with this pattern, a linear mixed effects model with random slopes for retrieval order and semantic/phonological similarity, and random intercepts for participants revealed a significant interaction between similarity and retrieval order ($b = .001$, $t = 4.19$, $p < .001$).
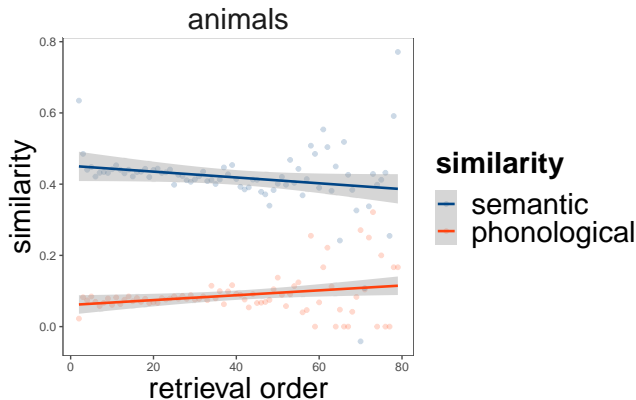

Figure 2: Semantic similarity decreased over time and phonological similarity increased over time in the fluency task.

Cues and retrieval performance

Next, we examined whether producing more semantically and/or phonologically related responses was related to the to-
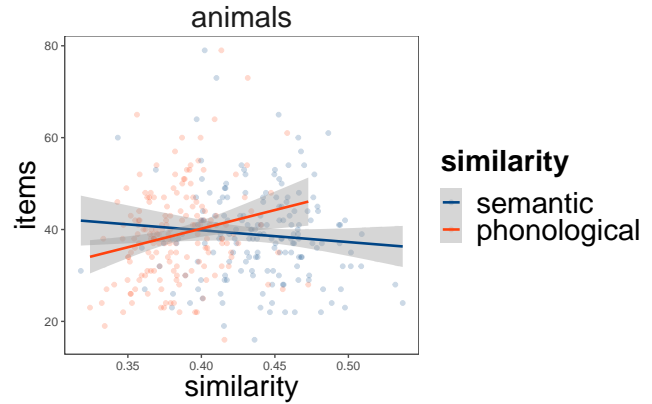

Figure 3: Number of items produced increased with phonological similarity and decreased with semantic similarity. Phonological similarity was scaled for plotting purposes.

tal number of items produced in the fluency task. As shown in Figure 3, greater mean semantic similarity (at the participant level) was associated with lower number of items produced, and greater mean phonological similarity was associated with greater number of items produced. This effect was confirmed by a significant interaction between similarity type (semantic/phonological) and items produced ($b = .001$, $t = 2.80$, $p = .006$) in a linear mixed effects model.

Clusters and switches

To better understand the role of phonological information in semantic search in SFT, we investigated whether semantic and phonological similarities across successive responses was correlated with switching designations defined in three ways. The first method corresponded to the Troyer (2000) norms, which contain hand-coded categorizations of animal terms (into subsets like "pets", "rodents" etc.). The second method corresponded to the similarity-drop model, as described in Hills et al. (2012), according to which switches are identified by noting where similarities drop between responses, in the following way: 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 third method was based on *participant-designated* switches in the LEA dataset, where participants indicated whether their responses were related in some way after the fluency task. Given that participant-designated switches were only available in the LEA dataset, we therefore restricted these analyses only to the LEA dataset. Figure 4 displays the mean phonological and semantic similarity as a function of switches and clusters across the three methods.

As shown in the top panel of Figure 4, greater semantic similarity was associated with "cluster" designations, compared to "switch" designations across all switch methods. However, greater phonological similarity was only associated with participant-designated "clusters" and not the similarity-
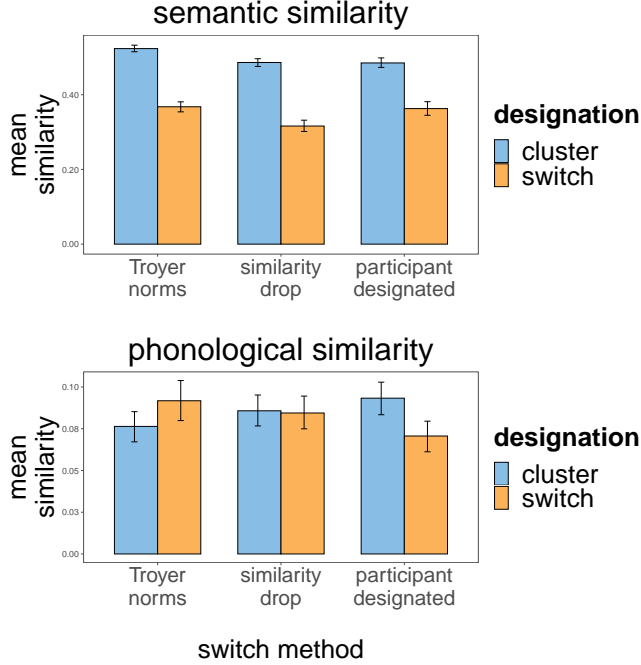
Figure 4: Participants grouped responses into clusters when semantic *and* phonological similarity was high, whereas the similarity-drop and Troyer-based models only identified clusters when semantic similarity was high. Error bars indicate bootstrapped 95% confidence intervals.

drop or Troyer (2000) clusters, as shown in the bottom panel of Figure 4. This pattern was confirmed by a significant three-way interaction between similarity type (semantic/phonological), switch designation method, and cluster/switch designations in a linear mixed effects model predicting mean similarity values ($b = .075$, $t = 4.97$, $p < .001$). Therefore, participants designated items into clusters based on *both* semantic and phonological similarity, whereas the similarity-drop and Troyer models did not capture this difference, due to being based on solely semantic similarity and hand-coded semantic categorizations.

Figure 5 displays some examples of response sequences classified as "clusters" by participants with high phonological similarity that were classified as "switches" by the similarity-drop and Troyer (2000) methods. As shown, although the *semantic* association decreased across the highlighted responses, *phonological* similarity increased, and participants perceived these responses to be related to each other.

Computational Models

Having established some key behavioral patterns showing an influence of phonology on semantic retrieval in the SFT, we next examined a series of computational search models that compared the relative influence of one, two, or all of three cues: semantics, phonology, and frequency on retrieval likelihoods of different responses produced in the fluency task. All models were derivatives of the optimal foraging model pro-
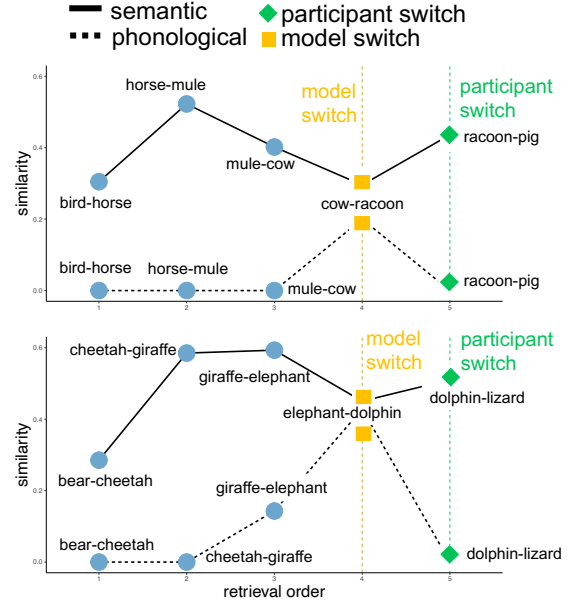

Figure 5: Response transitions classified as clusters by participants (cow→ racoon and elephant→ dolphin), that were classified as switches by the similarity-drop and Troyer models. While semantic similarity decreased from cow→ racoon and elephant→ dolphin, phonological similarity increased.

posed by Hills et al. (2012), which has been previously shown to capture search patterns in the SFT. In this model, the probability of retrieving an item *I* is computed as the product of the retrieval strength of that item relative to all other items produced in the sequence, described as follows:

$$P(I_i|q_1, q_2, ..q_M) = \frac{\prod_{j=1}^{M} S(q_j, I_i)^{\beta_j}}{\sum_{k=1}^{N} \prod_{j=1}^{M} S(q_j, I_k)^{\beta_j}} \quad (2)$$

where M is the total number of items produced, N is the size of the retrieval space, and β_j refers to the saliency of a given cue.

Frequency is a *global* cue, which generates the retrieval strength of an item based on its frequency in a natural language corpus. Semantic similarity is a *local* cue which generates the retrieval strength of an item based on its cosine similarity to the previous item in the response sequence. Within this framework, *dynamic* models exploited the patchy nature of the space, by selecting items using global information whenever there was a shift in the semantic subcategory (as defined by Troyer (2000) norms) or a drop in semantic similarity (as defined by the similarity-drop method). On the other hand, *static* models did not make use of any switch criterion to select the next item and simply used all cues to make transitions within semantic space. Phonological similarity was modeled as both a local and global cue across different models. Specifically, phonology was considered a local cue in the static models, but also considered as an additional cue for switching in the dynamic models, such that the next item

could be selected based on its phonological proximity to the previously retrieved item. Table 2 displays the 14 different models examined in the current work.[2]

Model comparisons

Using maximum likelihood estimation, we fit β parameters to each participant's data, for all cue types, using the participant's individually generated sequence of items. This produced a log-likelihood fit, which was penalized based on the number of free parameters in each model via the Bayesian information criterion (BIC). Results are presented as the median improvement in the Bayesian information criterion relative to a random model specifying that all remaining items in the search space are equally likely to be retrieved.

	type	within-category	switch
1	static	sem*freq	-
2		sem*freq*phon	-
3	Troyer-based	sem*freq	freq
4	dynamic	sem*freq	freq*phon
5		sem*freq*phon	freq
6		sem*freq*phon	freq*phon
7	similarity-drop	sem*freq	freq
8	dynamic	sem*freq	freq*phon
9		sem*freq*phon	freq
10		sem*freq*phon	freq*phon
11	participant-	sem*freq	freq
12	designated	sem*freq	freq*phon
13	dynamic	sem*freq*phon	freq
14		sem*freq*phon	freq*phon

Table 2: Search models evaluated in the fluency task

Table 3 displays the model fits for the different static and dynamic models fit to the fluency data based on whether the switch was made based on the Troyer norms, the similarity-drop method, or the participant-designated switches. First, consider the model comparisons reported for the static models and the Troyer-based dynamic models (models 1-6). As shown, the best-performing model among these was the dynamic model that used frequency, semantic, and phonological information to make within-category transitions and frequency for switches (i.e., model 5). This model showed the greatest improvement in model fit compared to a random model. Next, the similarity drop-based models generally provided better fits to the data compared to the Troyer-based models, consistent with Hills et al. (2012). Importantly, the model comparisons followed the same general trend, such that the model that used frequency, semantic, and phonological information to make within-category transitions and used frequency for switching (i.e., model 9) provided the best model fit over and above a random model.Finally, all the participant-designated switch models (models 11-14)

---

[2]Models 11-14 were only fit for the LEA dataset which contained the participant-designated switch values.

provided better fits than the static models, although here the model with phonology as a global cue led to a slightly higher improvement in model fit compared to the model with phonology as a local cue.

To evaluate whether participants benefited from the use of phonology, locally or globally, we performed a sign test of whether *any* model with phonology provided a better fit than the dynamic model without phonology. Phonology-based models provided better fits (lower BIC) for 103 of 171 participants based on the Troyer norms ($p = .005$), 99 of 171 participants based on the similarity drop method ($p = .023$), and 14 of 30 participants based on the participant designations ($p = .71$).

To summarize, we compared a series of foraging models with and without phonology, using three different methods to designate switches within the dynamic models. The static models did not designate any switches and used all cues locally to make transitions. Our model comparisons suggest that models that use semantic information, frequency *and* phonological information to make within-category transitions and switches provided better fits to the SFT data. Therefore, when individuals are searching for items within a given patch, they appear to use all three cues to retrieve the next item and find a new patch. However, given the marginal differences in model fit between the models that used phonology as a local or global cue, future work should explore the specific contribution of phonology in exploiting a given patch of items versus exploring new patches in memory.

Discussion

In this work, we examined the relative influence of semantic and phonological information on retrieval performance in the fluency task. A few findings are of note here. First, to our knowledge, this is the first study to examine the combined influence of semantic and phonological information on search in SFT. We found that while semantic similarity declined, phonological similarity increased over the course of the task. Next, an increase in phonological similarity was also associated with greater number of items retrieved. Furthermore, participants classified items that were phonologically similar as related after generating responses, whereas existing methods of capturing response transitions did not predict this pattern. Finally, we showed that a foraging model that combined semantics, frequency, and phonology outperformed models without phonology in predicting the retrieval likelihood of responses produced in the fluency task. We now discuss the implications of each of these findings.

The finding that semantic similarity decreases over the course of the fluency task is not surprising - a shorter response sequence (e.g., cat→dog) is likely to be more semantically similar than a longer response sequence (e.g., cat→dog→hamster→rhino). However, it is surprising that phonological similarity appears to increase over the duration of the task. One possibility is that as the task progresses, participants start to run out of semantically related items within

model	type	within-category	switch	Δ median BIC
1	static	sem*freq	-	56.45
2		sem*freq*phon	-	61.16
3	Troyer-based	sem*freq	freq	59.56
4	dynamic	sem*freq	freq*phon	61.70
5		**sem*freq*phon**	**freq**	**64.34**
6		sem*freq*phon	freq*phon	64.18
7	similarity-drop	sem*freq	freq	65.63
8	dynamic	sem*freq	freq*phon	66.57
9		**sem*freq*phon**	**freq**	**70.03**
10		sem*freq*phon	freq*phon	69.72
11	participant-designated	sem*freq	freq	67.34
12	dynamic	sem*freq	freq*phon	68.81
13		sem*freq*phon	freq	68.76
14		**sem*freq*phon**	**freq*phon**	**69.81**

Table 3: Median improvement in BIC for models fit to animal fluency data, compared to a random model; higher Δ values indicate better fit. Best-performing models have been highlighted in **bold**.

a neighborhood and alternative cues become more salient. Phonological similarity may be more salient when semantic similarity is no longer a useful cue in later stages of the task. This may also explain the second critical finding in these data: greater phonological similarity of responses was associated with greater number of items produced. It is possible that attending to this additional cue during semantic retrieval may help individuals capitalize on the multiplex nature of the lexicon (Stella et al., 2018) and ultimately aid retrieval (also see Levy et al., 2021; Castro et al., 2020; Castro & Stella, 2019). Our findings are consistent with this literature, and suggest semantic memory models may benefit from integration with other lexical information.

In addition to demonstrating the influence of phonology via behavioral measures, another important contribution of this work is the introduction of formal optimal foraging models that explore the combined influence of semantic, phonological, and frequency-based information on semantic retrieval processes. Specifically, our modeling analyses indicated that models that included phonological similarity as a within-category cue outperformed models without phonology. This result is consistent with the cluster-switch designations provided by participants (see Figure 4), where participants grouped semantically *and* phonologically similar items into clusters. Although there is the possibility that post-hoc participant designations may not reliably reflect the search processes occuring during the task, there is strong neural evidence to suggest that these participant-designated switches do indeed reflect active cognitive operations occurring in the brain during semantic search (see Lundin, in press; Lundin et al., in preparation). Together, these behavioral and computational modeling results suggest that phonology acts as a *local* cue that is used to produce items that are within the same semantic *and* phonological neighborhood.

On the other hand, the role of phonology in determining switch events during semantic search still needs to be fur-

ther investigated. Consider the response sequence aardvark→ ant→cat→ dog. Although it is clear that the transition from ant→cat was phonologically mediated, it also appears that this transition guided the participant to a new semantic cluster of *pets*, which led them to subsequently retrieve *dog* as their next item. It is possible that using only semantic information to designate switching events in our Troyer and similarity drop-based computational models may have obscured some of these effects. Specifically, existing methods of identifying clusters within fluency lists do not correlate well with participant-designated clusters. Indeed, in the LEA dataset, the correlation between participant and model-designated switches was only moderate for the Troyer ($r = .46$, $p < .001$) and similarity-drop models ($r = .35$, $p < .001$). Furthermore, models with phonology as a global cue performed better among the participant-designated models. It is possible that by implicitly accounting for phonology in within-category transitions by using participant-designated clusters, the model was better able to identify the role of phonology in global transitions. However, given that there were minimal differences across these models, the role of phonology in switching still needs to be further explored. Overall, however, the present results do provide some preliminary evidence that phonology may be an additional cue that determines how responses are clustered and when switches occur during semantic search. Therefore, future work should investigate alternative switching methods that can capture this source of variability in responses produced in the fluency task.

This work evaluated the role of phonology in retrieval from semantic memory in the SFT, and showed that individuals not only use phonology to guide their memory search, but also benefit from this additional retrieval cue. These findings suggest that computational models of semantic memory may need to account for these implicitly correlated cues (semantics and phonology) to better understand their interplay in cognitive tasks.

References

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, K. B., M J, Loftis, B., , & Treiman, R. (2007). The english lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, *113*(39), 10818–10823.

Castro, N., & Stella, M. (2019). The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks*, *7*(6), 913-931.

Castro, N., Stella, M., & Siew, C. S. (2020). Quantifying the interplay of semantics and phonology during failures of word retrieval by people with aphasia using a multiplex lexical network. *Cognitive Science*, *44*(9).

Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform similarity increases with semantic similarity: An analysis of 100 languages. *Cognitive Science*, *41*(8), 2149–2169.

Davelaar, E. J. (2015). Semantic search in the remote associates test. *Topics in Cognitive Science*, *7*(3), 494–512.

De Deyne, S., Navarro, D. J., Collell, G., & Perfors, A. (2019). Visual and affective multimodal models of word meaning in language and mind. *Cognitive Science*, *45*(1).

Dingemanse, M., Blasi, E, D., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in cognitive sciences*, *19*(10), 603–615.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431.

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.

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological review*, *114*(1), 1.

Kumar, A. A., Steyers, M., & Balota, A., David. (2021). A critical review of network-based and distributional approaches to semantic memory structure and processes topics in cognitive science. *Topics in Cognitive Science*, *45*. doi: https://doiorg/101111/tops12548

Lake, B. M., & Murphy, G. L. (2021). Word meaning in minds and machines. *Psychological Review*, *Advance Online Publication*. doi: https://doiorg/101037/rev0000297

Levy, O., Kenett, Y. N., Oxenberg, O., Castro, N., De Deyne, S., Vitevitch, M. S., & Havlin, S. (2021). Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. *Scientific Reports*, *11*(1), 1–14.

Lundin, N. B. (in press). Disorganized speech in psychosis: Computational and neural markers of semantic foraging and discourse cohesion. *Doctoral Dissertation, Indiana University, ProQuest Dissertations Theses Global*.

Lundin, N. B., Brown, J. W., Johns, B. T., Jones, M. N., Purcell, J. R., Hetrick, W. P., . . . Todd, P. M. (in preparation). Monitoring and control processes guide semantic and phonetic foraging in memory.

Lundin, N. B., Todd, P. M., Jones, M. N., Avery, J. E., O'Donnell, B. F., & Hetrick, W. P. (2020). Semantic search in psychosis: Modeling local exploitation and global exploration. *Schizophrenia Bulletin Open*, *1(1), sgaa011*.

Mikolov, T., Chen, G., Kand Corrado, & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:13013781*.

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*(3), 325.

Patel, A., Sands, A., Callison-Burch, C., & Apidianaki, M. (2018). Magnitude: A fast, efficient universal vector embedding utility package. *arXiv preprint arXiv:181011190*.

Stella, M., Beckage, N. M., Brede, M., & De Domenico, M. (2018). Multiplex model of mental lexicon reveals explosive learning in humans. *Scientific Reports*, *8*(1), 1-11.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, *22*(3), 370–378.

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in alzheimer's and parkinson's disease. *Journal of the International Neuropsychological Society*, *4*(2), 137–143.