Finding the right words: A computational model of cued lexical retrieval

Molly Apsel
Indiana University
mapsel@iu.edu

Abhilasha A. Kumar
Indiana University
kumaraa@iu.edu

Michael N. Jones
Indiana University
jonesmn@iu.edu

Abstract

Failing to come up with a word or name is a fairly common experience that is exacerbated in older adulthood and among populations with language impairments, and yet the mechanisms underlying lexical retrieval remain fairly understudied. In this work, we introduce and evaluate a series of nested computational models of lexical retrieval that combine semantic representations derived from a distributional semantic model with a process model to account for behavioral performance in a primed lexical retrieval task. The models were tested on a behavioral data set where participants attempted to retrieve answers to descriptions of low-frequency words and were provided a semantically and/or phonologically related prime word before the retrieval attempt. Model comparisons indicated that a model that emphasized semantic activations from the description and phonological activations from the prime word best accounted for the overall data. Additionally, incorrect responses and metacognitive judgments indicating that participants had other words in mind did not show this improved performance for models emphasizing phonological activations over semantic activations from the prime word. Taken together, these results identify the locus of lexical retrieval failures and offer the opportunity to investigate broader questions about semantic memory retrieval.

Keywords: lexical retrieval; semantic retrieval; semantic model; priming

Introduction

Everyone has had the frustrating experience of trying to come up with the name of a familiar person, place, or thing and being unable to access it in the moment. This inability to retrieve lexical items from memory can occur from failing to locate the correct semantic concept, or the phonological codes required for articulation, or both. Indeed, lexical retrieval failures are fairly common, tend to increase in older adulthood (Juncos-Rabadán et al., 2010; Logan & Balota, 2003), and are also used in diagnostic testing for semantic dementia (Calabria et al., 2021; Meteyard & Patterson, 2009) and aphasia (Friedmann et al., 2013; Herbert et al., 2008) in clinical settings. Therefore, it is important to understand the mechanisms underlying lexical retrieval.

An interesting aspect of failed lexical retrieval is that individuals tend to report partial access to information about the word, such as producing related associates, the number of syllables, or onset phonemes and often have the sense that the word is at the tip of their tongue (Brown & McNeill, 1966). Consequently, a large body of work has investigated lexical retrieval through behavioral manipulations, such as priming paradigms (Oberle & James, 2013; White et al., 2013) that manipulate the type of information available to individuals at the moment preceding retrieval. In these experiments, individuals typically attempt to retrieve words in response to a verbal or written description (e.g., the illegal act of writing untrue things about someone), and are primed with semantically (e.g., perjury) and/or phonologically (e.g., litigate/label) related information before the retrieval event. A robust finding in this literature is that phonologically related primes appear to facilitate lexical retrieval (James & Burke, 2000; Meyer & Bock, 1992), whereas semantic primes appear to have no such influence (Roediger & Neely, 1982; Kumar et al., 2019), suggesting that lexical retrieval failures arise due to impaired access to phonological information about the word. However, the explicit mechanisms by which this retrieval process is mediated continues to remain unclear. Specifically, empirical accounts of lexical retrieval tend to rely on the spreading activation metaphor (Collins & Loftus, 1975), and propose that activation "spreads" from the description and primes to different concepts in semantic memory, which ultimately leads individuals to retrieve the intended word. While these models describe the process of accessing a word from a semantic representation through spreading activation, they typically do not incorporate any computational model of semantic memory. Examining lexical retrieval within the context of an existing computational model of semantic memory can provide further insights into how concepts are accessed from memory during retrieval.

In this paper, we propose a computationally-driven account of lexical retrieval that leverages state-of-the-art machine learning-based methods for representing concepts in memory in conjunction with a process model to account for retrieval performance in a primed lexical retrieval task (see Kumar, 2021a for a similar approach)[1]. There is prior evidence that semantic and phonological information interact during lexical retrieval (Ferreira & Griffin, 2003), yet *how* these sources are combined in not well studied. Our approach formally instantiates and compares a series of computational models to identify how different sources of information (semantic and phonological) are combined to produce responses during retrieval from semantic memory. This modeling approach allows us to not only examine the conditions

[1] All data and analysis scripts are available at `https://github.com/abhilasha-kumar/modeling-lexical-retrieval`

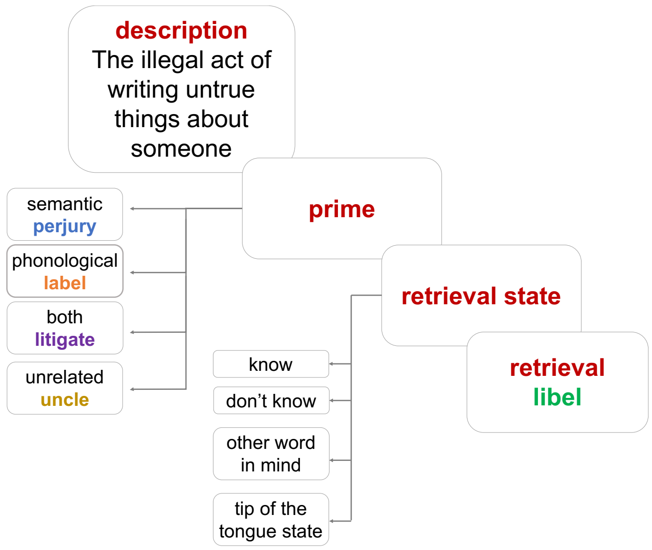

Figure 1: Experimental paradigm in Kumar et al. (2019).

under which successful lexical retrieval occurs, but also evaluate how incorrect responses are generated and discuss the broader implications for impaired lexical retrieval. While speech errors have been extensively studied in the lexical retrieval domain from a phonological perspective (Dell et al., 1997), we focus on how concepts systematically activated in *semantic* memory lead to retrieval errors. In this way, our approach provides a novel integration of a modern computational model of semantic memory with phonological information to predict lexical retrieval.

All models in the current paper assumed that lexical retrieval began by reading a description of the intended word and activating related concepts in semantic memory. These activations were then combined with semantic and/or phonological neighbors of other words in memory to ultimately produce a response. To test the predictions from our models, we used a publicly available data set of primed lexical retrieval performance collected by Kumar et al. (2019) and evaluated how different parameter settings within our model account for correct and incorrect responses, as well as phenomenological responses produced by participants in this data set.

Experimental task and data set

To test our model predictions, we utilized the publicly available lexical retrieval data set collected by Kumar et al. (2019). In this study, 174 participants read a description (i.e., the definition of a low-frequency target word), which was immediately followed by a "prime" word that was either phonologically related, semantically related, both phonologically and semantically related, or unrelated to the target word (see Figure 1). Before typing their answer to the description, participants also reported their metacognitive state from four options: (1) I know the answer; (2) I do not know the answer; (3) I have another word in mind that I don't think is correct; (4) The correct answer is "on the tip of my tongue". Each participant received 100 descriptions with one of the four prime types for each description during the experiment. The prime type presented in each trial was varied in random order and counterbalanced within participants.

Computational Models

We evaluated a series of nested computational models to account for lexical retrieval performance. Each model assumed the design of a standard primed lexical retrieval task, where participants are provided a description of the target word, presented a prime word, and then asked to retrieve the intended target word. Figure 2 displays the overall modeling framework, which involved the following steps:

1. A search space of 13,693 words was defined, which consisted of the 12,216 words produced by participants in a large free association data set (De Deyne et al., 2016) as well as all targets, primes, and valid responses in the Kumar et al. data set.[2]

2. A distributional semantic model, the Universal Sentence Encoder (USE; Cer et al., 2018) was used to obtain vector representations of all words in our search space and the 100 descriptions in the data set. Our choice of the USE for studying lexical retrieval was motivated by the recent success of neural network-based language models in capturing long-range contextual dependencies in natural language. Previous work in this domain has focused on associative accounts of semantic memory, which are based on behavioral norms (see Kumar et al., 2021 for a discussion). Distributional models directly encode statistical regularities in natural language and therefore circumvent the circularity issue. The USE model goes a step further and incorporates information about word order and linguistic context during the creation of semantic representations for longer sequences, which allows for meaning to vary as a function of context and provides a computational account for how phrases or longer sequences can be represented. This is consistent with modern context-driven accounts of semantic memory (Kumar, 2021b; Yee & Thompson-Schill, 2016) and therefore represents a novel integration of state-of-the-art models of semantic memory with theories of lexical retrieval. We used the Deep Averaging Network (DAN) variant of the Universal Sentence Encoder, which passes averaged vector representations from words and bigrams to a deep neural network and is trained on multiple language tasks. The DAN encodes text of any length into 512-dimensional vectors. These vector representations can then be used to estimate similarity between concepts within a given semantic space. Note that this distributional search space is isomorphic to a fully-connected weighted semantic network, where each word is connected to every other word in the vocabulary based on its semantic similarity.

[2] 14 responses were excluded due to being invalid words such as "omnipo," "neoprology", and "obstrecian"

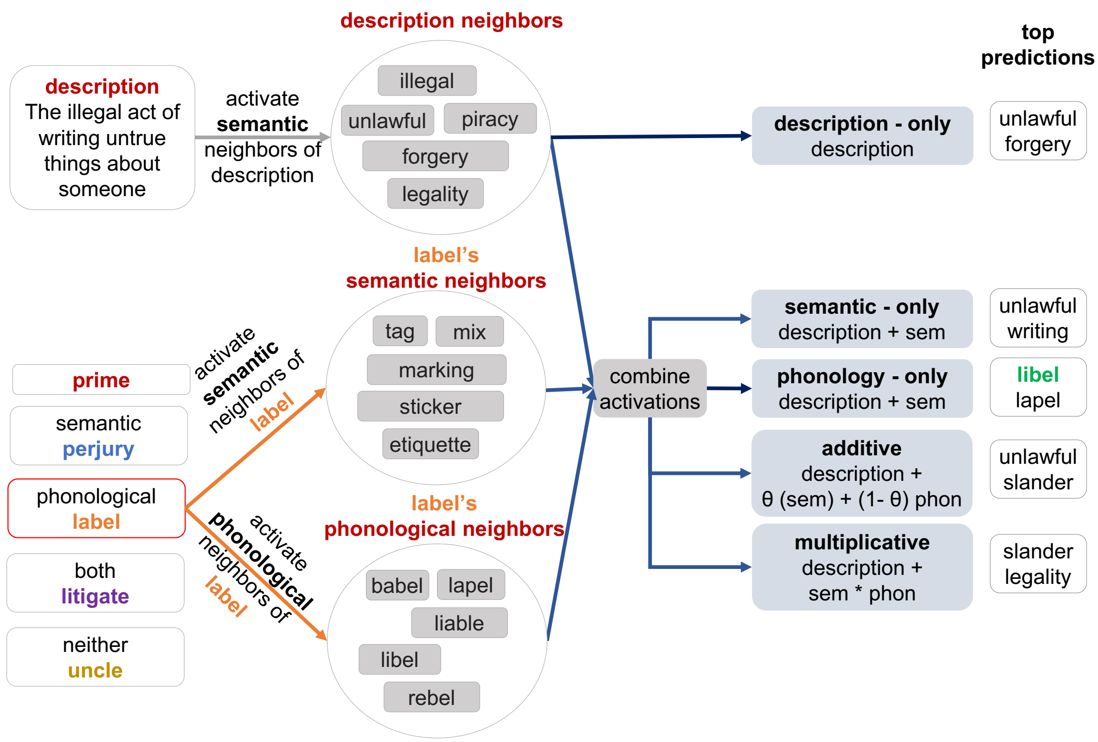

Figure 2: Modeling framework for the primed lexical retrieval task. The description and the prime (label) independently activate words in the search space, which are then combined to produce retrieval likelihoods for different responses. In this example, the phonology-only model retrieves the correct answer to the description for the target word, libel.

3. Each word in the vocabulary had an activation value at all times, representing the activation of that word relative to all other words in the search space. Each trial started with setting the initial activations of all words to zero.

4. The **description-only** model began by reading the description and activating all its neighbors in proportion to their cosine similarity with the description. The resulting values represented the semantic activation of each word given the description, such that the description's closest neighbors had the highest activations. Therefore, this model evaluated the likelihood of a given word in the search space after processing only the description.

5. The next set of models evaluated the contribution of the prime to the retrieval process by emphasizing different aspects of the prime word:

   - The **semantic-only** model activated the prime's semantic neighbors in the same way as the description-only model. Therefore, the resulting activation in this model was the sum of activations obtained from the description as well as the prime.
   - The **phonology-only** model activated the prime's phonological neighbors by computing a measure of phonological similarity between the prime and each word in the search space. Phonological similarity was computed as the normalized edit distance between the phonemes contained within the words. Phonemes were obtained using the CMU Dictionary, which relies on the Arpabet phonetic transcription. Normalized edit distance was computed as follows:

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

     where e(a,b) signifies the edit distance between *a* and *b*, and l_a and l_b denote the lengths of *a* and *b*.
   - The **additive** and **multiplicative** models explored the combined influence of the prime's semantic and phonological activations. Additive models used element-wise addition to combine the two prime components, varying the weight given to each one with a parameter θ, as follows:

$$a_i \propto d_i + \theta(s_i) + (1-\theta)(p_i) \tag{2}$$

     where a_i denoted the activation of a given word *i*, d_i denoted the activation of i from the description-only model, s_i denoted the activation of i by semantic neighbors of the prime, and p_i denoted the activation of i by phonological neighbors of the prime. θ varied from .1 to 0.9 across different additive models. The multiplicative model combined the semantic and phonological activations via element-wise multiplication:

$$a_i \propto d_i + s_i p_i \tag{3}$$

description	prime	model	predictions
Unconventional and slightly strange; deviating from an established or usual pattern or style	ecstatic	desc-only sem-only phon-only multiplicative additive (θ=.5)	unorthodox thrilled eccentric eclectic erratic
To goad or push forward; to incite someone to do something bad	initiate	desc-only sem-only phon-only multiplicative additive (θ=.5)	provoke encourage infuriate motivate instigate

Table 1: Example predictions from the different models evaluated on the lexical retrieval task.

In all models, the final resulting activations corresponded to the relative likelihood that a given word in the search space was the answer to the description. Table 1 displays a few example predictions from each of the models for different descriptions and prime combinations[3].

Model comparisons

We evaluated the performance of the different models (description-only, semantic-only, phonology-only, multiplicative, and the nine θ- based additive models) based on how well they predicted behavioral responses in the Kumar et al. data set, by computing response likelihoods under the different models.

Trials on which the participant gave no response (43.32% of the total responses) were excluded. We computed summed log likelihoods (LL) for all the models and then evaluated models based on the Bayesian Information Criterion (BIC). BIC penalizes models with more parameters; therefore, the semantic-only and phonology-only models were penalized for having 1 additional parameter compared to the description-only model (the prime's semantic or phonological activations), the multiplicative model was penalized for using both semantic and phonological activations (2 additional parameters), and the additive models were penalized for having 1 more parameter compared to the multiplicative model (θ).

Overall model performance Table 2 displays the overall performance of the different models, ordered by BIC. As shown, the description-only model had the worst model fit (highest BIC). Model performance improved progressively across the additive models as θ (weight on semantics) decreased, and the phonology-only model provided the best fit to the overall data (lowest BIC).

Prime-based performance Given that the type of prime provided was a critical manipulation in the Kumar et al. (2019) data set, we next evaluated the success of the models in retrieving the target across different prime conditions. Figure 3 displays the BIC for the different models across the four prime conditions. For *semantic* primes, the semantic-only model provided the best fit. Given that the semantic primes were not phonologically related to the target word, any weight given to the phonology of the prime increased activation of incorrect words and decreased the likelihood of activating the target. For *phonological* primes, the phonology-only model best predicted the responses. Thus, any weight given to the semantics of the prime lured the model away from the description's semantic neighbors, increasing the likelihood of incorrect answers and decreasing the likelihood of the target. Model performance for *both* primes showed a slightly muted but similar pattern to the phonological primes, such that the phonology-only model provided the best fit. Finally, for unrelated primes, the phonology-only model also performed best, but the multiplicative and description-only models performed better than most other models. Therefore, in situations when the semantic and phonological components of the prime word did not provide any information relevant to the target, participants were about as likely to be influenced by phonological cues as they were to use the multiplicative model or ignore the prime altogether.

model	parameters	BIC
description-only	1	182923.00
multiplicative	3	180680.93
semantic-only	2	180376.00
additive	4	
θ = 0.9		180283.23
θ = 0.8		180193.68
θ = 0.7		180107.36
θ = 0.6		180024.28
θ = 0.5		179944.48
θ = 0.4		179867.97
θ = 0.3		179794.77
θ = 0.2		179724.90
θ = 0.1		179658.39
phonology-only	**2**	**179595.26**

Table 2: Overall model performance, reported as BIC. The **phonology-only** model best fit the behavioral data.

Correct and incorrect responses Although the analyses on the full data set are useful, it is important to recognize that the responses produced by participants may correspond to correctly retrieving the target in some cases, and producing incorrect words in other cases. Therefore, we analyzed model differences separately for correct and incorrect responses. Figure 4 displays the model performance for correct and incorrect responses. As shown, correct responses mirrored the pattern of overall responses, such that the phonology-only model provided the best account of the data. On the other hand, incorrect responses were about equally well explained by all of the additive models. Therefore, when participants

[3] Table 1 only displays *one* additive model but nine additive models corresponding to different θ values were evaluated

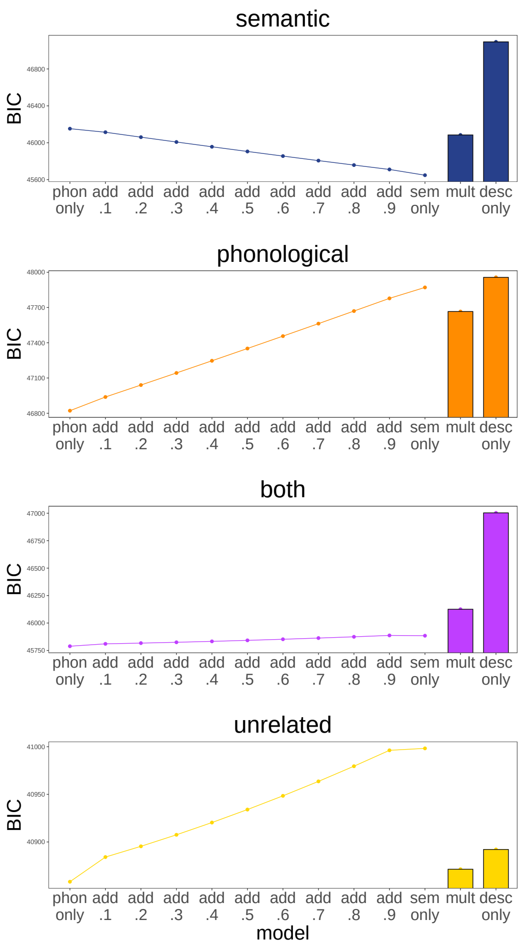

Figure 3: BIC indices for models fit on the lexical retrieval data for different prime conditions; lower indices indicate a better model fit.

successfully utilized a combination of the semantic activation from the prompt and the phonological activation from the prime, they were more likely to retrieve the correct answer.

Retrieval states In addition to evaluating the retrieval likelihood of different responses, we also explored whether the models could differentiate between the metacognitive reports provided by participants. Table 3 shows the performances of the models for the different retrieval states indicated by participants. As shown, "know", "don't know", and "tip-of-the-tongue" responses were best explained by the phonology-only model, whereas the "other word in mind" responses were best explained by the semantic-only model. When the data were broken down by both prime type and retrieval state, the pattern was similar to Figure 3, with one exception. While trials preceded by *both* primes were generally better accounted for by the phonology-only model, when participants chose "other word in mind", their responses were better predicted by the semantic-only model. Figure 5 displays the number of incorrect responses, correct responses,

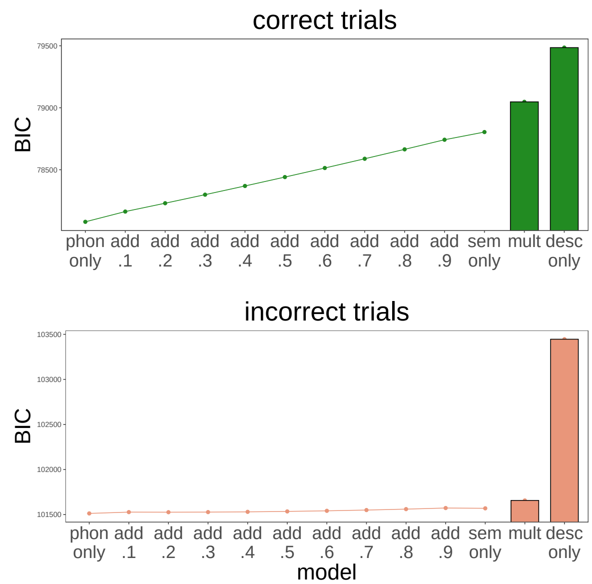

Figure 4: BIC indices for correct and incorrect trials.

and trials with no response for each of the self-reported states. As shown, when participants reported having another word in mind, they were indeed lured by the semantic information contained in the prime word and were most likely to type these "other" incorrect responses instead of the target word. This pattern was best captured by the semantic-only model. On the other hand, the proportion of incorrect responses was lower for other retrieval states, which were better explained by the phonology-only model.

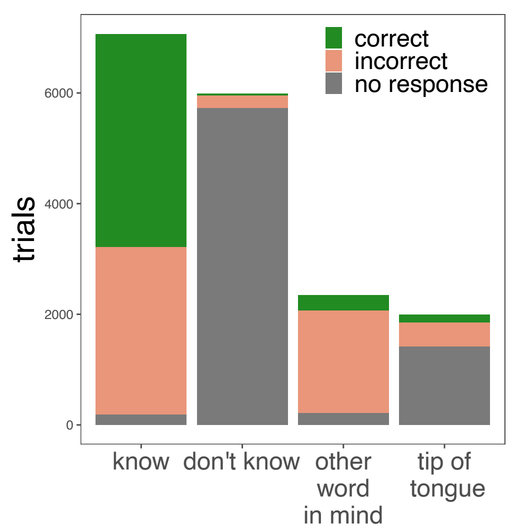

Figure 5: Proportion of correct responses, incorrect responses, and no-responses for each reported state.

model	parameters	BIC			
		know	don't know	other word	tip of the tongue
description-only	1	127539.89	4835.04	39808.58	10758.70
multiplicative	3	125878.46	4789.09	39401.74	10669.30
semantic-only	2	125676.10	4779.48	**39292.27**	10648.19
additive (mean of all 9 models)	4	125301.54	4783.05	39307.57	10640.15
phonology-only	2	**124933.68**	**4766.01**	39354.37	**10610.26**

Table 3: Model BIC by retrieval state. Best-fitting model for each retrieval state has been highlighted in each column.

Discussion

In the present research, we implemented and evaluated a series of nested computational models to elucidate the mechanisms that best explain successful and unsuccessful lexical retrieval. To our knowledge, this is the first computational instantiation of a lexical retrieval model that combines a state-of-the-art machine learning model of semantic representation with a process model of how these representations are accessed and retrieved within a given task. We tested the models on a large behavioral data set of primed lexical retrieval, where participants were given a description of a low-frequency target word, followed by a prime word that related to the target in one of four ways (phonologically, semantically, both, or unrelated), and attempted to retrieve the intended target word. We measured the likelihood of each model to best account for the responses generated by participants. We found that a model that combined the semantic activations from the description with phonological activations from the prime word provided the best account of the data. These results converge with prior work (Kumar et al., 2019; James & Burke, 2000; Meyer & Bock, 1992) suggesting that phonological cues facilitate lexical retrieval.

In addition to evaluating overall model performance, a second novel contribution of this work is that we were able to examine the likelihood of different responses, when exposed to different types of primes, when retrieval was successful and unsuccessful, as well as across different metacognitive states reported by participants. Previous work on lexical retrieval has typically focused on retrieval accuracy as the sole dependent measure. However, our findings suggest that examining the specific responses generated by participants provides deeper insights into the mechanisms underlying lexical retrieval. For example, when breaking down the data by prime type, models that emphasized the prime's semantic information best predicted responses when participants were primed with semantic information. Furthermore, when the prime contained relevant phonological information (i.e., in the case of phonological and "both" primes), phonology was more valuable than semantic information because it provided unique information about the target word not contained within the description. Interestingly, even when the prime was completely unrelated to the target, the semantic-only model performed worse than the phonology-only model, likely because it diluted the semantic activations from the description. Overall, these findings suggest that activating the semantic *and* phonological codes of a word is critical to successful lexical retrieval.

Another interesting finding in our model comparisons was that correct responses reflected a clear use of prime phonology over prime semantics while incorrect responses did not clearly reflect one use of prime cues over another. Furthermore, when participants reported having another word in mind, their responses showed a greater reliance on semantic information from the prime. The vast majority of such cases also produced an incorrect response, suggesting that their attention to the prime's semantic component led them to activate the wrong semantic space. These comparisons shed light on the situations that might lead a person to retrieve incorrect words, and suggest that attending to the phonological information may be more beneficial than attending to semantic information in the moment before retrieval. Of course, in everyday lexical retrieval, individuals do not have access to primes to facilitate or inhibit their retrieval process. Therefore, an important future step for this work is to evaluate the parameters that best account for unprimed lexical retrieval performance.

Finally, the current approach only allowed us to examine situations when individuals explicitly produced a response. Yet, it is equally important to analyze the cases in which people fail to retrieve any word, which made up nearly half of trials in the Kumar et al. (2019) data set. In future work, we plan to implement an activation threshold for different responses within our modeling framework to simulate situations in which no response is generated. This would allow us to more carefully examine the conditions that produce a failed lexical retrieval event, which may in turn have broader applications for aging and clinical populations with impaired semantic retrieval. Another goal is to incorporate further degrees of spreading activation into our models. In the present instantiation, semantics and phonology simultaneously activate words for retrieval. However, prior theories of lexical selection describe activation that continuously spreads bidirectionally between concepts, words, and phonemes. In future models, we hope to include these features of spreading activation to more fully account for the process of lexical retrieval in semantic memory.

References

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337. doi: http://dx.doi.org/10.1016/S0022-5371(66)80040-3

Calabria, M., Jefferies, E., Sala, I., Morenas-Rodríguez, E., Illán-Gala, I., Montal, V., … Costa, A. (2021). Multilingualism in semantic dementia: Language-dependent lexical retrieval from degraded conceptual representations. *Aphasiology*, *35*(2), 240–266.

Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., John, R. S., … others (2018). Universal sentence encoder. *arXiv preprint arXiv:1803.11175*.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and non-aphasic speakers. *Psychological review*, *104*(4), 801.

Ferreira, V. S., & Griffin, Z. M. (2003). Phonological influences on lexical (mis) selection. *Psychological Science*, *14*(1), 86–90.

Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and its breakdown in aphasia and developmental language impairment. *The Cambridge handbook of biolinguistics*, 350–374.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology*, *22*(2), 184–203.

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1378.

Juncos-Rabadán, O., Facal, D., Rodríguez, M. S., & Pereiro, A. X. (2010). Lexical knowledge and lexical retrieval in ageing: Insights from a tip-of-the-tongue (tot) study. *Language and Cognitive Processes*, *25*(10), 1301–1334.

Kumar, A. A. (2021a). Modeling semantic structure and spreading activation in retrieval tasks [doctoral dissertation]. washington university in st. louis. *ProQuest Dissertations Publishing*.

Kumar, A. A. (2021b). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review*, *28*(1), 40–80.

Kumar, A. A., Balota, D. A., Habbert, J., Scaltritti, M., & Maddox, G. B. (2019). Converging semantic and phonological information in lexical retrieval and selection in young and older adults. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *45*(12), 2267–2289.

Kumar, A. A., Steyvers, M., & Balota, D. A. (2021). A critical review of network-based and distributional approaches to semantic memory structure and processes. *Topics in Cognitive Science*.

Logan, J. M., & Balota, D. A. (2003). Conscious and unconscious lexical retrieval blocking in younger and older adults. *Psychology and Aging*, *18*(3), 537-–550. doi: https://doi.org/10.1037/0882-7974.18.3.537

Meteyard, L., & Patterson, K. (2009). The relation between content and structure in language production: An analysis of speech errors in semantic dementia. *Brain and Language*, *110*(3), 121–134.

Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory Cognition*, *20*(6), 715–726.

Oberle, S., & James, L. E. (2013). Semantically-and phonologically-related primes improve name retrieval in young and older adults. *Language and cognitive processes*, *28*(9), 1378–1393.

Roediger, H. L., & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *36*(2), 213.

White, K. K., Abrams, L., & Frame, E. A. (2013). Semantic category moderates phonological priming of proper name retrieval during tip-of-the-tongue states. *Language and Cognitive Processes*, *28*(4), 561–576.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic bulletin & review*, *23*(4), 1015–1027.