Do Taxonomic and Associative Relations Affect Word Production in the Same Way?

Delaney C. McDonagh (dmcdonag@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Ave. Pittsburgh, PA 15213 USA

Anna V. Fisher (fisher49@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Ave. Pittsburgh, PA 15213 USA

Nazbanou Nozari (bnozari@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Ave. Pittsburgh, PA 15213 USA

Abstract

Naming a picture is more difficult in the context of a taxonomically-related picture. Disagreement exists on whether non-taxonomic relations, e.g., associations, have similar or different effects on picture naming. Past work has reported facilitation, interference and null results but with inconsistent methodologies. We paired the same target word (e.g., cow) with unrelated (pen), taxonomically-related (bear), and associatively-related (milk) items in different blocks, as participants repeatedly named one of the two pictures in randomized order. Significant interference was uncovered for the same target item in the taxonomic vs. unrelated and associative blocks. There was no robust evidence of interference in the associative blocks. If anything, evidence suggested that associatively-related items marginally *facilitated* production. This finding suggests that taxonomic and associative relations have different effects on picture naming and has implications for theoretical models of lexical selection and, more generally, for the computations involved in mapping semantic features to lexical items.

Keywords: word production; semantic interference, semantic facilitation, taxonomic similarity, associative similarity

Introduction

Semantic interference, defined as slower and/or more error-prone production of words in the context of semantically related words, has been a robust finding in picture naming studies using both continuous naming methods and repeated naming of a small set of items (e.g., Schnur et al., 2006). The majority of such studies have used taxonomically related items, i.e., items that were members of the same category, e.g., "cow" and "bear", which are both animals. The presence of robust semantic interference has long been taken as support for an influential theoretical account of lexical selection: the *competitive selection account* poses that production cannot proceed until a certain differential threshold has been reached between the activation of the target lexical item and competitors. Because of their shared features with the target, semantically related lexical items are particularly prone to gaining activation, and semantically related context exaggerates this. Thus, a general prediction of the competitive selection account was that semantically related context should create interference in production.

This prediction was violated when it was discovered that non-taxonomic relations do not yield the same interference effect in picture naming. Picture-word interference paradigms have often shown that naming a picture is in fact *facilitated* by presenting a distractor word that is associated with the picture in a non-categorical manner. For example, associates facilitate the production of other associates (e.g., bone → dog; Alario, Segui & Ferrand, 2000). Based on this evidence, the *non-competitive selection view* was formed, which suggested that semantic relations, in general, facilitate lexical selection, and the interference arises at other levels (e.g., Mahon et al., 2007). Recently, a third view has been proposed, suggesting that production can have a differential threshold, as proposed by the competitive selection account, but this threshold is flexible and can be changed based on task goals and the state of the production system to mimic a non-competitive profile (Nozari & Hepner, 2019). This account, however, views the change in the threshold as a process separate from the dynamics of spreading activation in the semantic-lexical system. The question thus remains: Do different kinds of semantic relations have different effects on picture naming?

To explain the discrepancy between the interference created by taxonomic similarity on the one hand, and the facilitation effect created by non-taxonomic relations in picture interference word paradigms on the other hand, Abdel Rahman and Melinger (2009) proposed the "swinging lexical network" account. In a nutshell, they proposed that semantic similarity induces both semantic facilitation and lexical interference. Taxonomically related items induce interference because they activate a cohort of lexical items via a common category node (e.g., bear → cow, horse, etc., via the *animals* node). Associative items are not connected via a category node, and thus do not activate additional cohort. In short, the differential effect is ascribed to the number of activated competitors (many vs. one) in taxonomic vs. associative relations; both types of relations induce some facilitation, but only the former activates enough lexical items to create enough interference to counter the facilitation.

Abdel Rahman, Melinger and colleagues tested the predictions of the swinging lexical network model using a cyclic blocked naming paradigm, in which participants repeatedly named a small set of items in taxonomically related, associatively related, and unrelated blocks. The logic was that, compared to picture word interference paradigms, repeated naming of multiple items should lead to interference even for associatively related items. This was indeed what was found (Abdel Rahman & Melinger, 2007; 2011; Aristei, Melinger & Abdel Rahman, 2011). De Zubicaray et al. (2014) criticized these studies by pointing out that the associative relations were contaminated by taxonomic features. When they redesigned the study with materials that were better controlled in this respect, no evidence of interference was uncovered in the associative condition. Instead, de Zubicaray et al. (2014) found that such relations generated a transient facilitation effect compared to the unrelated condition. This finding casts doubt on the explanation offered by the swinging lexical network account. More generally, it remains unclear whether different types of semantic similarity have fundamentally different effects on word production, because at least two main issues have been overlooked in the designs of prior studies.

The more prominent of these issues is the absence of a definition of "associative" relations, beyond a general description of occurring in a common setting or theme (hence the alternative label, "thematic"), without belonging to the same category. It is difficult, however, to measure the strength of such associations. For example, the associative category "United States" in Aristei et al. (2011) contains items like "prairie", "hamburger", and "cap". But in the absence of explicitly linking them together with a label "United States", these items are not highly associated in the language network. In keeping with this, some of the interference effects found for associative relations were only present when participants were explicitly presented with a verbal label (Abdel Rahman & Melinger, 2011). Similarly, de Zubicaray et al.'s (2014) materials include blocks such as "Roman" with members such as "lion" and "shield" which are not otherwise strongly associated, which may be the reason for their null effect. The second issue is that none of these studies have controlled for phonological similarity which is documented to induce interference effects (e.g., Nozari et al., 2016). It is thus possible that some of the interference stems from sources other than the hypothesized one. The current study aimed to investigate the effect of taxonomic and associative relations on picture naming in a design that avoided these problems.

The current study

The current study used a shorter version of the cyclic blocked picture naming task (Nozari et al., 2016). Participants saw only two pictures per block and named them repeatedly, one at a time, in randomized order for a total of 12 trials per block. The same targets (e.g., "cow") were paired with three "competitor"[1] pictures in the three conditions: a taxonomically related item (e.g., "bear"), an associatively related item (e.g., "milk"), and an unrelated item (e.g., "pen"). The design has the following features:

(a) It still builds on the strength of the cyclic naming paradigm and avoids extra processes such as reading or hearing irrelevant distractors. It is thus more likely to reflect the true dynamics of the production system. Moreover, the repeated naming of the two pictures maximizes the activation in the semantic-lexical network and helps us investigate the trajectory of the effect beyond initial naming attempts. This is particularly important given the reports of transient facilitatory effects on first trials in such a paradigm (e.g., Schnur et al., 2006). At the same time, the shortened version avoids the variability induced by differences in working memory demand.

(b) The main difference between the current and prior studies is quantifying "associative similarity". Items that frequently occur together in the world, also occur frequently together in language. In fact, modern techniques of analysis in the field of distributional semantics have shown that the similarity structure in the real world can be faithfully recovered from the structure of language (e.g., Mikolov et al., 2013). The problem is that many items are related to other items in more than one way. For example, "cat" and "dog" are members of the same animal category, but also frequently appear in the same real-life scenarios, as well as in language. Thus, in order to test pure effects of taxonomy and association, one needs 1) objective measures for each, and 2) pairs that show a double dissociation on such measures, i.e., score high on one and low on the other and vice versa. This rationale was applied to the current study.

To quantify taxonomic similarity, we used Resnik scores (Resnik, 1995), calculated on word pairs in WordNet. WordNet is an implementation of the mental lexicon and organizes four classes of words (nouns, verbs, adjectives, and adverbs) with regard to their semantic relationships to other words. The key relationship that forms the organization of WordNet is hyponymy/hypernymy, or superordinate relations, which generates a hierarchical network in which various semantic relations such as synonymy, meronymy, and antonymy are represented. For any given word pair, a Resnik similarity score represents their closeness in these taxonomic hierarchies, with 0 indicating no relationship and higher scores indicating greater similarity.

To quantity associative similarity, we used pointwise mutual information (PMI), which indexes the probability that the two words cooccur in text, with PMI = 0 indexing co-occurrence at chance level and positive and negative values indicating more and less than chance probabilities, respectively. Results were double-checked with another measure, log-likelihood (LL), since the two measures have been shown to have different biases (Evert, 2007).

Using these measures, we created a stimulus set in which targets were paired with competitor items that showed a

---

[1] From the perspective of the participant, there is no difference between targets and competitors. All pictures are named.

double dissociation in their taxonomic and associative relations to the target. For example, for the taxonomic condition, target "cow" was paired with "bear". This pair had a high Resnik and low PMI and LL scores. For the associative condition, target "cow" was paired with "milk". This pair had high PMI and LL and low Resnik scores. The unrelated competitor (e.g., "pen") was chosen to have low scores on both taxonomic and associative measures.

(c) The competitor items were matched carefully with each other and with target items in length, lexical frequency, and age of acquisition. Phonological overlap was tightly controlled and kept close to zero in all sets. Finally, (d) we selected highly familiar, high-frequency items, with age of acquisition of all items <6 years. This choice, together with the selection of items with high PMI and LL scores in the associative condition, is important, because it is virtually impossible for adult, native speakers who name the two pictures in the associative conditions repeatedly not to notice the common theme, and thus removes the concern that the manipulation may have been too weak for associative effects to show up.

These characteristics make the design well suited for comparing the effect of taxonomic vs. associative similarity in picture naming. We first compare the effect on the same (target) word, which is identical in all three conditions. Since the materials are controlled for other psycholinguistic factors, we would expect the effect to be replicated when the competitor items are also included. All analyses are based on the same predictions: if different kinds of semantic similarity have the same general effect on word production, we should see consistent effects across taxonomic and associative conditions. On the other hand, qualitatively different patterns of data would point to fundamentally different dynamics resulting from the two kinds of similarity.

Methods

Participants

Forty-three native English-speaking undergraduate students (M_{age}= 19.33, *SD*=1.80; 24 female) participated for credit.

Materials

Three conditions (taxonomic, associative, and unrelated) were created, each containing six pairs of monosyllabic words that were related taxonomically (e.g. *cow/bear*), or associatively (e.g. *cow/milk*), or were unrelated (e.g. *cow/pen*). Targets were the same, while the competitor words were different in each set, creating a total of 24 unique words (Table 1). Items were matched for lexical frequency, age of acquisition, and word length. The sets were also matched across condition for phonological similarity which was minimized by design. As explained above, taxonomic and associative competitors had high Resnik and PMI/LL measures respectively, while scoring low on the other measure. Resnik scores were calculated on WordNet, as $\text{Sim}_{\text{Resnik}}$ (c1, c2) = IC(LCS(c1, c2)), where LCS(c1, c2) = Lowest node in hierarchy that is a hypernym of c1, c2; lowest common subsumer, and IC(c) = -logP(c); information content. LL scores were calculated with Natural Language Toolkit for Python, using a window of 5 words (and replicated with a window of 10 for robustness) on SubtlexUS (Brysbaert & New, 2009).

Taxonomically related pairs were chosen to have high Resnik scores (*M* = 5.69), and low associative scores (PMI: *M* = .79; LL: *M* = .69). Conversely, associatively related pairs had high associative scores (PMI: *M* = 5.20; LL: *M* = 32.23), and low Resnik scores (*M* = 1.18). Unrelated pairs had low scores on both sets of measures (Resnik: *M* = 1.62; PMI: *M*= .51; LL: *M* = 0.196).

Twenty-four 400 x 400-pixel, color images corresponding to the 24 words were selected from Google Images, taking care to minimize visual similarity (e.g., for animal pairs, postures were selected that did not always reveal four legs, etc.). For cases where some similarity in the taxonomic condition was unavoidable due to category membership (*cow/bear*), we included a pair that was more visually similar in the associative condition (*hand/glove*).

Three pseudo-randomized orders were created, where the same item was not repeated more than three times in a row within a block and the same target never appeared in adjacent blocks. The orders were counterbalanced across conditions to ensure an equal number of repeats and switches across conditions, with roughly twice as many switches as repeats, creating three different lists. Each list contained 18 blocks (6 of each condition), each containing one pair of pictures.

Table 1. The six sets used in the experiment. Each row shows the target words, along with the three competitors. T = taxonomic, A = associative, U = unrelated.

Set	Target	T	A	U
1	Cow	Bear	Milk	Pen
2	Car	Bike	Road	Doll
3	Hand	Tongue	Glove	Chair
4	Sock	Hat	Foot	Kite
5	Bird	Sheep	Nest	Plate
6	Pear	Corn	Tree	Boat

Procedure

The experiment was run in MATLAB R2019a with the Psychtoolbox-3 package. Pictures were displayed at the center of a 20 x 13inch Dell desktop approximately 25 inches in front of the participants. Participants were assigned to one of three lists (see above). On each block, they first saw both pictures and read their written labels aloud. Next, they completed four practice trials. On each trial, one of the two pictures was presented for 1500 ms (or until a response was recorded) and participants were instructed to name the picture as quickly and accurately as possible. The next trial started after 500 ms of a blank screen. After practice, participants completed 18 blocks of 12 experimental trials (6 of each picture), with the same structure as practice trials. Response times (RTs) for spoken responses were registered using an Audio-Technica microphone, and all spoken responses were recorded.

Coding and Analyses

All data were transcribed and coded offline for accuracy, which, as expected, was near ceiling in neurotypical adult speakers. All the RTs registered with the voice key were double checked by viewing the acoustic wave for individual words in PRAAT and measuring the onset of the word from the onset of the beep (which marked the onset of the picture). Word durations were also calculated by marking the offset of words in PRAAT using text grids, with manual checking, and subtracting the onsets.

Analyses were conducted in R (version 3.6.1; R Core Team, 2019). Trials with incorrect responses and RTs more extreme than 3 standard deviations away from each participant's mean RTs were excluded from the analyses. For the analyses, RTs were log-transformed to better approximate a normal distribution. Data were analyzed by linear mixed effects models (LMEMs) using lme4 (version 1.1-21), in conjunction with lmerTest (version 3.1-0) to calculate p-values using Satterthwaite approximations. Comparisons of interest were contrast-coded in the models as taxonomic vs. unrelated, associative vs. unrelated, and taxonomic vs. associative. We included the maximum random effect structure that the models tolerated, which was the random intercepts for participants and items.

Results

In total, 9288 responses were collected (4644 on target items). Error rates were low (<1%) and did not differ between conditions. After excluding errors and extreme RTs, ~3% of data were excluded from analysis. RTs were 453 (SE=7.8), 441 (SE=7.5) and 444 (SE=7.7) ms in taxonomic, associative and unrelated conditions, respectively. The results of the LMEMs showed that RTs were significantly longer in the taxonomic vs. both unrelated ($\beta = -.02$, $t = -3.51$, $p < 0.001$), and associative ($\beta = -.03$, $t = -4.62$, $p < 0.001$) conditions. While the RTs were numerically shorter in the associative vs. unrelated condition, this difference did not reach significance ($\beta = .01$, $t = 1.11$, $p = .27$). A very similar pattern was observed for the whole dataset (targets + competitors; Figure 1a). The models' results were also very similar; significantly longer RTs for the taxonomic vs. unrelated ($\beta = -.02$, $t = -3.54$, $p < .001$) and associative ($\beta = -.03$, $t = -4.66$, $p < .001$) conditions, and no significant effect of associative vs. unrelated condition ($\beta = .007$, $t = 1.14$, $p = .26$). Figure 1b shows the target RTs for the three conditions, graphed by trials that contained targets in the randomization (all trials except for trials 6 and 11). As seen in the graph (and expected from the past literature), the effect of the first trial is very different than the rest of the trials.

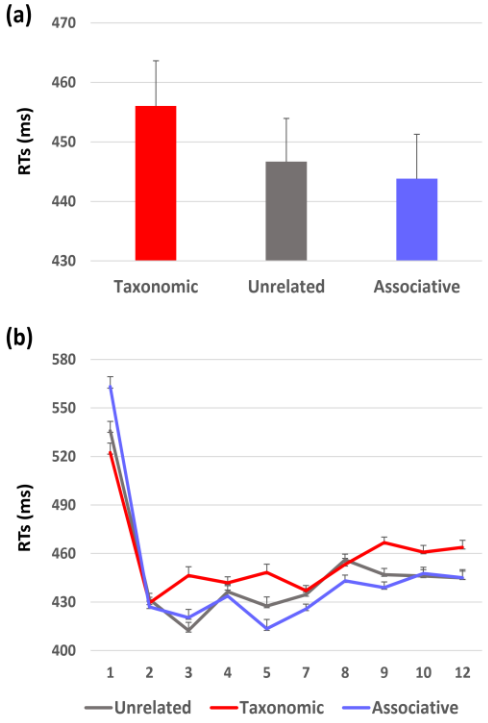

Figure 1: RTs in the three conditions. (a) Average RTs by condition for all items. The pattern is identical to the Target-only dataset. Error bars show by-participant SEs. (b) Target RTs in the three conditions by trial. Trials 6 and 11 are not shown because those positions did not contain target items in any of the three randomized lists. The other positions contain an equal number of target items in the three conditions. Error bars are by-trial SEs.

The taxonomic effect shows the same pattern as previous reports from blocked cyclic naming paradigms with a larger set: a temporary facilitation followed by interference. The associative condition shows a flipped pattern. Because of the clearly different status of the first trial, we ran the same model excluding the first trial[2]. The effect of taxonomic vs. unrelated and associated conditions were unchanged ($\beta = -.022$, $t = -3.53$, $p < 0.001$, and $\beta = -.029$, $t = -4.62$, $p < 0.001$). But this time, there was also a marginal effect of associated vs. unrelated condition ($\beta = .01$, $t = 1.8$, $p = .07$). The same analyses repeated with the whole set returned very similar results: significantly slower RTs for taxonomic vs. unrelated and associative conditions ($\beta = -.025$, $t = -3.58$, $p < .001$; and $\beta = -.029$, $t = -4.66$, $p < 0.001$, respectively), and marginally faster RTs in the associated vs. unrelated condition ($\beta = .011$, $t = 1.8$, $p = .07$). Importantly, note that this marginal effect was in the direction of *facilitation*, i.e., the opposite of the effect of taxonomic condition. Finally, to ensure that we were not missing a tradeoff between RTs and durations, durations were also compared using similar models with the same contrasts. There were no significant differences in durations between the taxonomic (343 ms; SE=37) and unrelated (344

---

[2] We also ran a model with an interaction between condition and trial, with trial coded in a binary fashion, as "first" vs. "others". The problem is the few number of trials in the first compared to other trials, which creates an imbalance in the analysis. Nevertheless, the analysis showed no interaction between condition and trial for the contrast coding taxonomic vs. unrelated but revealed a significant interaction for the contrast coding associated vs. unrelated, which is in line with the findings of the analysis reported in the text.

ms; SE = 37) conditions (β =.079, t = .38, p = .71), between taxonomic and associative (345 ms; SE = 37) conditions (β = 1.56, t = .91, p = .36), or between associative and unrelated conditions (β = -1.19, t = -.57 , p = .57).

Discussion

This is, to our knowledge, the first study to test the effect of semantic and associative similarity in picture naming using materials that were selected based on quantitative measures to show a double dissociation in their relationship to the target. This is important for two reasons: first, it ensures that the hypothesized relationship (taxonomic or associative) is indeed strong, which, as pointed out earlier, has been a problem in the past studies, making the interpretation of null effects (e.g., de Zubicaray et al., 2014) particularly difficult. Second, a double dissociation is necessary to ensure that the uncovered effects of specific kinds of semantic similarity are not cross-contaminated. Using this method, and controlled for additional factors, we replicated the interference effect induced by taxonomic similarity in an alternative version of the cyclic blocked naming task with only two items. However, we found no evidence that associative relations affected the target production in the same way as taxonomic relations. More specifically, not only was there no robust evidence of interference in this condition, but also the effect was in the opposite (facilitatory) direction. In fact, save the first trial, our data showed a marginally significant facilitation induced by an associated competitor. Moreover, direct comparisons between taxonomic and associative conditions consistently revealed a significant difference. All of these effects were robust on targets which remained constant across conditions, as well as the whole dataset which also included comparisons between different but well-controlled competitors across conditions.

These findings, together with the absence of any significant differences in accuracy and word durations, show unambiguously that taxonomic and associative contexts affect word production quite differently. The findings are better aligned with previous results that associations, if anything, tend to facilitate production (e.g., Alario et al., 2000; de Zubicaray et al., 2014) and show that that the absence of interference cannot be attributed to the lack of repetition. The argument that participants may not have been aware of the relationship between words in the associative blocks is also extremely unlikely, given the highly transparent relations (e.g., cow/milk, car/road, hand/glove, etc.). Finally, proponents of the swinging lexical network might suggest that the effect is due to the small set, as they claim that taxonomic relations activate an array of cohorts that belong to the category but associative relations do not. The rationale for this claim, as well as the empirical evidence in its support, is unclear. If anything, the long-standing concepts of “schema”, i.e., quick activation of a large set of elements related to a common scenario is based purely on associations (e.g., Rumelhart et al., 1986). It is thus hard to argue that repeated visualization and naming of pictures of cow/milk activated a weaker schema (and hence fewer items) compared to cow/bear.

We must also mention one previous study that investigated the effects of non-taxonomic similarity, specifically avoided lexical associations: de Zubicaray et al. (2014) stated that the materials were selected based on “informal polling” of an independent group of participants and that, according to free association norms, they were *not* associates. We would like to re-emphasize that semantic and lexical associations are highly correlated, as shown by the success of distributional semantics approaches that use linguistic regularities to uncover structured semantic knowledge (e.g., Mikolov et al., 2013). It is thus not only impractical to try to separate semantic and lexical association, but more importantly, doing so will inevitably lead to relations that are semantically weak; perhaps too weak to generate any reliable effects.

In summary, the results of this study strongly suggest that taxonomic and associative relations induce different effects in picture naming.

Theoretical Significance

As noted in the Introduction, the interference effect induced by various kinds of semantic relationships has been taken as evidence for competitive lexical selection. The fact that two words that are highly lexically associated (and thus highly activate each other) do not interfere with one another, does not lend support to this view. Note that, if anything, lexical associations activate lexical competitors a lot more strongly than any other kind of semantic similarity, so if interference really results from close *lexical* activations per se, repeated production of two lexical associates should create the largest interference. The current data do not support this prediction.

An alternative account of semantic interference, the incremental learning account, offers an explanation without resorting to competitive lexical selection. For example, Oppenheim, Dell, and Schwartz’s (2010) model explains interference in the following manner: naming an item leads to the strengthening of the connections between the semantic features of that lexical item (e.g., four legs and “cow”), and weakening of the same features to other items that share those features (e.g., “bear”). This differential weight change puts those other items at a relative disadvantage if they become the target on the next trial. This manifests as semantic interference. Thus, the behavioral finding of semantic interference is not equivalent to a competitive lexical selection mechanism. Another possibility is that strict selection is not really mandatory at the lexical level, and that the system can proceed with a good deal of cascaded activation and narrow down selection at lower levels.

But the finding that associative relations have a fundamentally different effect on picture naming than taxonomic relations is still theoretically relevant to the predictions of models such as Oppenheim et al.’s. Most computational models of word production are agnostic about the nature of semantic representations. In such models, semantic similarity is represented by a subset of semantic nodes that are shared between the lexical nodes of taxonomically similar items. The simplest (and seemingly reasonable) way to model associative relations in such models is to do something similar; connect a number of associatively similar lexical nodes to a shared set of semantic nodes, arguing that while the *nature* of the semantic overlap is different, the *dynamics* of activation and incremental weight adjustments should be the same. The current results clearly oppose this view and suggest that to best understand the cognitive processes underlying word production, one must pay careful attention to the organization of semantic knowledge and how that knowledge interacts with lexical representations.

Preliminary proposals in this regard include a superordinate node for the taxonomically related words (e.g., "animals" for cow/bear), which does not have an equivalent in associatively related words (Aristei et al., 2011). But it is hard to theoretically refute that a "theme" node cannot play a similar role (e.g., "dairy farm" for cow/milk). It is also hard to argue that the supraordinate node automatically activates a lot more coordinates than the theme node, given the concept of schema discussed earlier. Another possibility is that the weight changes act in a more selective manner to maximize production success according to task demands. Although the majority of studies that investigate semantic interference control for visual similarity as much as possible, the fact remains that taxonomically related items, by definition, share "features", most of which create visual confusion in general, as well as in the context of specific tasks. Such a close interaction between visual features and the language production system is absent in associative relations. It is thus possible that the system has learned to maximize differences when similarity affects the input system, at least in the context of a visually driven task such as picture naming.

While these explanations are speculative, the bottom line is clear: different kinds of semantic similarity affect word production in different ways. This finding, in turn, suggests that the two kinds of semantic similarity cannot be modeled the same way in computational models of word production. One solution could be to directly connect lexical nodes of associated items. This may require revising the assumption of competitive selection at the lexical level. A second solution would be to acknowledge that the nature of semantic knowledge is critical to the dynamics of weight change in the language production system. Consequently, the dynamics of incremental learning must be modeled with regard to the nature of shared knowledge, and in accordance with what maximizes production efficiency in a goal-oriented manner. These solutions are not mutually exclusive. Future modeling work should explore the ramifications of either choice.

Acknowledgements

This work was supported by the NSF grant BCS-1949631 to N.N., and the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition-Scholar Award (220020401) to A.V.F.

References

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*(3), 604.

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effect in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes, 24*(5), 713-734.

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 149.

Alario, F.X., Segui, J., Ferrand, L. (2000). Semantic and Associative Priming in Picture Naming. *The Quarterly Journal of Experimental Psychology, 53*(3), 741-764.

Aristei, S., Melinger, A., & Abdel Rahman, R. (2011). Electrophysical Chronometry of Semantic Context Effects in Language Production. *Journal of Cognitive Neuroscience, 23,* 1567-1586.

Brysbaert, M. & New, B. (2009) Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behavior Research Methods*, *41*(4), 977-990.

de Zubicaray, G., Johnson, K., Howard, D., & McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex, 54*, 135-149.

Evert, S. (2007). Corpora and collocations. *Corpus linguistics. An international handbook*, *2*, 1212-1248.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503-535.

Mikolov, T., Sutskever, I., Chen, K. Corrado, G. S., Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems.*

Nozari, N., Freund, M., Breining, B., Rapp, B., & Gordon, B. (2016). Cognitive control during selection and repair in word production. *Language, Cognition and Neuroscience*, *31*(7), 886–903.

Nozari, N., & Hepner, C. R. (2019). To select or to wait? The importance of criterion setting in debates of competitive lexical selection. *Cognitive Neuropsychology*, *36*(5-6), 193-207.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*(2), 227-252.

Resnik, P. (1995). Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In Mellish, C.S. (Ed.), *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 1, 448-453.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In D. E. Rumelhart and J. L. McClelland, (Eds.), Parallel distributed processing: explorations in the microstructure, vol. 2. Cambridge, MA: MIT press.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199-227.