



Published in final edited form as:

*Cognition*. 2010 February ; 114(2): 227–252. doi:10.1016/j.cognition.2009.09.007.

## The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production

Gary M. Oppenheim<sup>a</sup>, Gary S. Dell<sup>a</sup>, and Myrna F. Schwartz<sup>b</sup>

<sup>a</sup> Beckman Institute, University of Illinois, 405 N. Matthews Ave., Urbana, IL 61801, USA

<sup>b</sup> Moss Rehabilitation Research Institute, 1200 West Tabor Road, Philadelphia, PA 19141, USA

### Abstract

Naming a picture of a dog primes the subsequent naming of a picture of a dog (repetition priming) and interferes with the subsequent naming of a picture of a cat (semantic interference). Behavioral studies suggest that these effects derive from persistent changes in the way that words are activated and selected for production, and some have claimed that the findings are only understandable by positing a competitive mechanism for lexical selection. We present a simple model of lexical retrieval in speech production that applies error-driven learning to its lexical activation network. This model naturally produces repetition priming and semantic interference effects. It predicts the major findings from several published experiments, demonstrating that these effects may arise from incremental learning. Furthermore, analysis of the model suggests that competition during lexical selection is not necessary for semantic interference if the learning process is itself competitive.

### Introduction

Retrieving a word from memory has consequences for later retrieval. This is particularly true when retrieval occurs in a semantic memory task such as picture naming. It is well known that the second presentation of a picture to be named speeds the naming response and diminishes the chance of error. This phenomenon, known as *repetition priming*, can be explained by the fact that each retrieval event is also a learning event, and so the second retrieval benefits from the learning that occurred the first time (*e.g.* Mitchell & Brown, 1988). Somewhat less well known is the fact that repetition priming has a “dark side”. Retrieving a word has *negative* consequences for the subsequent retrieval of other words from the same semantic category (*e.g.* Belke, 2008; Belke, Meyer, & Damian, 2005; Blaxton & Neely, 1983; Brown, 1981; Damian & Als, 2005; Damian, Vigliocco, & Levelt, 2001; Howard, Nickels, Coltheart, Cole-Virtue, 2006; Hsiao, Schwartz, Schnur, & Dell, 2009; Kroll & Stewart, 1994; Schnur, Schwartz, Brecher, & Hodgson, 2006; Vigliocco, Vinson, Damian, & Levelt, 2002; Wheeldon & Monsell, 1994). Following Oppenheim, Dell, and Schwartz (2007), we refer to these negative consequences as *cumulative semantic interference*. In this paper, we explain the mechanisms behind cumulative semantic

Address for correspondence: Gary M. Oppenheim, Beckman Institute, University of Illinois, 405 North Matthews Avenue, Urbana, Illinois 61801, goppenh2@illinois.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

interference in the domain of picture naming. This explanation takes the form of a computational model of lexical access in speech production that simulates the major phenomena in this domain. The model addresses meaning-based lexical retrieval in general, whether this is elicited by picture-naming, naming-to-definition, or spontaneous production. Our focus, however, is on persistent changes to lexical processing that result from the natural retrieval of a single word. The central theoretical point that the model implements is that repetition priming and cumulative semantic interference are two sides of the same coin. They both result from an error-based implicit learning process that tunes the language production system to recent experience.

Although our model is formally developed only for lexical access in speech production, our theoretical goals are more general. Cumulative semantic interference is a manifestation in speech production of a set of phenomena known in the memory literature as *retrieval-induced forgetting* or RIF. Retrieval-induced forgetting studies demonstrate that the episodic memory for a word or association can be impaired by the previous retrieval of a related memory (e.g. Anderson, Bjork, & Bjork, 1994; but see also Anderson & Neely, 1996, for a discussion of retrieval-induced forgetting in semantic memory). Currently, the explanation for such impairment is debated, with some claiming it results from suppressing previous competitors (often termed *inhibition* or *unlearning*; e.g. Anderson *et al.*, 1994; Melton & Irwin, 1940, Postman, Stark, Fraser, 1968) while others claim it stems from strengthening previous targets (*occlusion* or *'blocking'*<sup>1</sup>; e.g. McGeoch 1942; Mensink & Raaijmakers, 1988; MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003). Our analysis of cumulative semantic interference in speech production will, we claim, speak to this debate. More generally, our model reflects a recent trend in cognition to link psycholinguistics with theories of learning and memory by developing accounts of how experience changes language processing (e.g. Chang, Dell, & Bock, 2006; Goldinger, 1998; Kraljic & Samuel, 2005).

Much of the theoretical importance of cumulative semantic interference hinges on an alleged property of requiring a competitive mechanism for lexical selection (e.g. Howard *et al.*, 2006). The most prominent theories of lexical access (e.g. Levelt, Roelofs, & Meyer, 1999) assume competitive lexical selection. Empirical support for this assumption has often come from picture-word interference studies (e.g. Schriefers, Meyer, & Levelt, 1990), in which speakers name pictures as they are presented at short offsets from distractor words. However, since Mahon, Costa, Peterson, Vargas, and Caramazza (2007) presented an analysis demonstrating that picture-word interference studies have not reliably supported the claims of competitive lexical selection, the search for empirical support has turned to a simpler task: picture naming, specifically with regards to cumulative semantic interference.

Two serial picture-naming paradigms have been particularly common in studies of cumulative semantic interference. First is the *blocked-cyclic naming paradigm* (e.g. Damian *et al.*, 2001). In each block, subjects repeatedly cycle through naming a small set of pictures (e.g. one block might consist of four cycles through a set of six pictures). In the homogeneous condition, all the pictures in the block represent the same semantic category (e.g. farm animals), and in the mixed condition each picture represents a different semantic category. Cumulative semantic interference is indexed by greater difficulty naming pictures in the homogeneous condition relative to the mixed condition (the *semantic blocking effect*). Typically, the semantic blocking effect is not present in the first cycle and grows over subsequent cycles (e.g. Belke *et al.*, 2005). The second important serial picture-naming paradigm, used by Brown (1981, Experiment 4) and Howard *et al.* (2006), can be called the

---

<sup>1</sup>The term 'blocking' carries a quite different meaning in the retrieval-induced forgetting literature, where it refers to a hypothesis of competitor-based interference, than in the cumulative semantic interference literature, where it tends to refer to the structure of an experimental design (i.e. pictures may be presented in semantically homogeneous blocks).

*continuous paradigm*. In this method, pictures drawn from several categories (*e.g.* animals, vehicles) are named without repeating any item, but with multiple exemplars from each category. Here, cumulative semantic interference is demonstrated by naming times that increase linearly as a function of the number of previously named pictures in that category. Importantly, the number of interspersed pictures between each category exemplar is irrelevant to the effect (Howard *et al.*, 2006). For example, in the sequence *goat, car, tomato, truck, horse*, naming time for *horse* would be slower than that for *goat*, and the response times would be unaffected by the number of unrelated intervening items.

### The nature of cumulative semantic interference: Howard et al.'s principles

Howard *et al.* (2006) argued that three specific properties of the lexical retrieval process must interact to produce cumulative semantic interference in naming latencies: shared activation, competitive selection, and priming. The idea is that each time a target word is activated, semantically related competitors are also activated (*shared activation*), and strongly activated competitors slow down the selection of target words (*competitive selection*). Retrieving a word once primes its future retrieval (*priming*), making it a stronger competitor when related words are retrieved in subsequent trials, thereby causing those subsequent target words to be retrieved more slowly. We will use these three properties to structure our review of the phenomenon and its implications for lexical retrieval.

**Shared activation**—When a target word such as DOG is activated during its attempted retrieval, its semantic relatives such as GOAT are also activated, thereby setting the scene for lexical competition. This principle of *shared activation* for semantically related words is what makes cumulative semantic interference specifically *semantic* in nature.

While the idea of shared activation is compatible with most current theories of semantic representation, it arises naturally from the use of distributed (or feature-based) semantic representations such as those commonly employed in connectionist models (see McClelland & Rogers, 2003, for a review). Distributed mechanisms would predict graded effects of semantic similarity, and indeed blocked-cyclic picture naming studies have demonstrated that more closely related items generate stronger interference effects than those more distant (Vigliocco *et al.*, 2002). So, for the purpose of understanding cumulative semantic interference, it may be useful to think of shared activation arising from shared semantic features rather than all-or-none category membership. That is how shared activation is implemented in our model.

As noted by Howard *et al.* (2006), however, shared semantic activation does not require distributed representations. It may occur with non-decomposed (localist) lexical concepts (*e.g.* Roelofs, 1992) provided that related concepts connect either directly (*e.g.* Collins & Loftus, 1975) or indirectly through shared category or property nodes (*e.g.* Collins & Quillian, 1969), and each activated concept sends activation to neighboring concepts. Moreover, any graded effects can be attributed to gradations in the number or strength of such connections. Thus, the finding of graded cumulative semantic interference does not allow us to distinguish between distributed and localist semantic representations.

**Competitive selection**—The second property of lexical retrieval that is required for cumulative semantic interference, according to Howard *et al.* (2006), is that lexical selection be *competitive*. That is, increasing the activation of non-target words should decrease the speed and accuracy with which a target word is selected. In a competitive selection process, words compete in the manner of two athletic teams during “sudden-death overtime”: the competition continues until a single winner emerges. This might be implemented via either a differential threshold (*e.g.* Levelt *et al.*, 1999) or lateral inhibition (*e.g.* Howard *et al.*, 2006),

but the key is that having multiple strong competitors makes it harder to select a winner (Wheeldon & Monsell, 1994). A non-competitive selection process (*e.g.* Mahon *et al.*, 2007), in contrast, is more like a horse race that ends when the first contestant crosses a predetermined absolute threshold. To illustrate this difference, let us imagine selecting a target word, DOG, when a competitor, GOAT, is also activated. According to a sudden-death competition method, the two compete until one clearly wins, so selecting DOG should be slower and less accurate when GOAT is more active. Thus, cumulative semantic interference in response time would occur if the semantic manipulations raise the activation of competitors. With a horse-race selection method, the speed of DOG's selection is entirely a function of DOG's own activation. The activation of GOAT does not enter into the equation, so this non-competitive selection offers no obvious way to account for cumulative semantic interference.

Locating the competition within the word production process is difficult, but several studies constrain it to a point after semantic access and before phonological access. Two findings argue for a post-semantic locus. First, performing non-verbal semantic judgments on pictures in the blocked-cyclic paradigm has proven insufficient to elicit a semantic blocking effect (Damian *et al.*, 2001). So any competition that occurs during stages before lexical access does not appear sufficient to drive the cumulative semantic interference effect. Second, bilingual continuous paradigm experiments indicate that cumulative semantic interference accumulates independently for each language, suggesting that the competitive selection process is language-specific and hence post-semantic (Castro, Strijkers, Costa, & Alario, 2008, Experiments 3 & 4).

Some evidence suggests that the competition may instead characterize the selection of abstract, pre-phonological word forms, or lemmas. Blocked-cyclic *word* naming (reading words aloud) appears to produce semantic facilitation rather than interference, suggesting that the competition that results when naming pictures must arise before retrieving phonological word-forms (Damian *et al.*, 2001, Experiment 2a). Retrieving gender-marked determiners during word naming may bring back the semantic blocking effect, suggesting that the competition affects the post-semantic, pre-phonological retrieval of abstract lexical concepts (Damian *et al.*, 2001, Experiment 2b).

Together, the principles of shared lexical activation and competitive lexical selection are sufficient to produce the sort of semantic interference that might be seen within a single trial (*e.g.* as in the picture-word interference effect, *e.g.* Schriefers *et al.*, 1990). Shared activation causes semantically related competitors to become active, and a competitive lexical selection mechanism allows these competitors to hinder selection of a target word. Making semantic interference cumulative, however, requires some mechanism by which processes during one trial can affect subsequent trials. This is the function of priming.

**Priming**—Priming is Howard *et al.*'s final necessary property for cumulative semantic interference. Retrieving a word once should facilitate its future retrieval by either making the word itself more accessible or making its competitors less accessible.

While priming can be implemented in a number of ways, its effects can be characterized as either temporary or persistent. Temporary effects occur when priming is ascribed to changes in activation levels—either positive (*e.g.* Crowther, Martin, & Biegler, 2008; Howard *et al.*, 2006; Wheeldon & Monsell, 1994) or negative (inhibitory) changes (*e.g.* Brown, 1981; McCarthy & Kartsounis, 2000) that are carried over from previous trials. For example, selecting DOG may require the temporary suppression of the urge to say CAT, which might make it more difficult to access CAT for a short time. Persistent accounts (*e.g.* Damian & Als, 2005; Howard *et al.*, 2006; Schnur *et al.*, 2006) instead describe priming as a

consequence of relatively permanent changes to the way words are accessed, such as incremental learning. Inspired in part by neural network models in which incremental learning is attributed to changes in connection weights rather than activation levels, persistent priming is an example of the learning that continually adjusts the cognitive system to suit its environment (*e.g.* Gupta & Cohen, 2002).

A critical property of the priming mechanism that underlies cumulative semantic interference is that the interference accumulates incrementally as a function of relevant experience (such as naming semantically related pictures), and is unaffected by irrelevant experience (such as naming unrelated pictures). In Howard *et al.*'s (2006) continuous paradigm study, naming pictures from a single category, such as DOG and then GOAT, produced the same linear accumulation of semantic interference whether the relevant pictures were separated by two, four, six, or eight unrelated items, suggesting that only the relevant experience matters. Moreover, in a variant of Howard *et al.*'s (2006) continuous paradigm, Navarrete, Mahon, and Caramazza (2008) showed that repeating an item, such as DOG, GOAT, DOG produced the same cumulative interference as accessing additional novel exemplars from a category, thus demonstrating that each act of retrieval contributes to the effect.

Further evidence of robustness to irrelevant experience comes from the blocked-cyclic paradigm. Damian and Als (2005) showed that performing nonlinguistic tasks (Experiment 1) and naming unrelated items (Experiments 2 & 3) in between naming DOG and GOAT failed to disrupt the semantic blocking effect. Thus, the priming from each relevant experience contributes separately and robustly to the cumulative semantic interference effect, and irrelevant experience affects neither its accumulation nor diminution.

Filler trials, as used in the Howard *et al.* (2006) and Damian and Als (2005) studies required additional time to present and process, increasing the chronological time between the retrieval of related words. So these studies speak to residual activation (or inhibition) accounts of priming. Priming by residual activation should be strongly affected by the time between prime(s) and target. As argued by Bock and Griffin (2000), the activation levels that control language production must decay quickly in order for production — the rapid sequential activation of linguistic units — to succeed. For example, computer simulations of multi-word production have required that activation levels decay with time constants such that activated linguistic units lose nearly all of their activation within a second or two (*e.g.* Dell, 1986). Similarly, the effects of inhibitory processes in production are also time-bound. For example, many production theories assume that selection of a linguistic unit entails a reduction to a zero or negative activation value for the selected unit (*e.g.* Dell, Burger & Svec, 1997; Houghton, 1990). However, the effects of this inhibition are quite temporary and are designed to prevent an immediate perseveratory error. The fact that the filler trials in Howard *et al.* (2006), and Damian and Als (2005), failed to affect cumulative semantic interference suggests that the priming that underlies this effect is reasonably persistent. Hence, cumulative semantic interference therefore is likely not largely based on the positive or negative changes in activation levels that arise solely through the spreading activation mechanisms of the production system.

A more direct demonstration of the temporal insensitivity of priming comes from Experiment 1 of Schnur *et al.* (2006), who compared naming latencies in a blocked-cyclic naming paradigm in which pictures were presented either one second or five seconds after the previous response (*i.e.* a 1s or 5s response-stimulus interval, or RSI). Any time-based decay of residual activation predicts a statistical interaction between presentation rate and semantic blocking condition, specifically less effect of the blocking with the long RSI (*e.g.*, Wilshire & McCarthy, 2002). Both presentation rates produced reliable cumulative semantic

interference effects, with no interaction between RSI and semantic blocking condition, demonstrating that the priming is insensitive to the passage of time, at least at these intervals.

To summarize, the priming that causes cumulative semantic interference is temporally persistent, it accumulates with relevant experience, and it is insensitive to irrelevant experience. These properties offer an awkward fit for mechanisms based on residual activation or inhibition of linguistic units, both of whose effects should be expected to decay rather quickly. Instead, we follow Damian and Als (2005), Schnur *et al.* (2006), and Howard *et al.* (2006) by suggesting that the priming that underlies cumulative semantic interference emerges from small, persistent, experience-driven, post-selection adjustments to the mapping from semantics to words, *i.e.* incremental learning within the production system.

### Speech Errors and Cumulative Semantic Interference

It appears that cumulative semantic interference does more than slow lexical retrieval; it also causes lexical selection errors. In a blocked-cyclic naming task with healthy older controls, Schnur *et al.* (2006, Experiment 1) found that naming latencies were higher and errors more frequent in the homogeneous condition, relative to the mixed condition. They also tested aphasic patients (Schnur *et al.*, 2006, Experiment 2), who made many more errors, and reported two important findings. First, patients made more semantic errors (*e.g.* naming DOG as CAT) and omissions in the homogeneous than mixed condition. Second, these semantic blocking effects increased across cycles, while other types of errors (*e.g.* phonological) showed the opposite pattern. The patients' increasing semantic blocking effects for semantic and omission errors thus resembled healthy adults' increasing blocking effects for naming latencies, suggesting that they might stem from the same underlying causes.

The link between the blocking effects on errors in patients and on latencies in unimpaired speakers also has some support from studies that attempt to associate these effects with brain regions. Neuroimaging of healthy subjects demonstrated that activation in the left inferior frontal gyrus (LIFG) correlates with increases in naming latencies due to semantic blocking and related manipulations (Moss *et al.*, 2005; Schnur *et al.*, 2009). The LIFG, for reference, corresponds to Brodmann's areas (BA) 44, 45, and 47, the posterior part of which (BA 44/45) is Broca's area. And lesion analyses of patients from the Schnur *et al.*, (2006) study revealed an association between LIFG damage and the increase in errors across blocking cycles (Schnur, Lee, Coslett, Schwartz, & Thompson-Schill, 2005; Schnur *et al.*, 2009).

A second important finding from these patient studies is that patients' error effects are also robust to timing manipulations. Schnur *et al.* (2006) found that the blocking effect on errors – like the blocking effect on naming latency with unimpaired speakers – was not influenced by whether pictures were named with a 1 sec or a 5 sec RSI. Further examining these patients' naming errors, Hsiao *et al.* (2009) found that their within-set perseverations tended to match the words that they had used most recently. For instance, if a patient named pictures of a dog, a pig, and a goat correctly (*i.e.* saying PIG DOG GOAT) before incorrectly naming a picture of a horse, then she was more likely to name the horse as DOG than as PIG. Crucially, the key measure of recency was not time, but the number of intervening items (henceforth *item-lag*). Specifically, the chance-corrected perseveration lag functions were the same regardless of whether pictures were presented one second or five seconds after a patient's most recent response. Together with the temporal insensitivity of unimpaired speakers' response time effects, these results support our previous suggestion that relevant intervening experience, not timing, matters for the build-up or dissipation of cumulative semantic interference.

## Modeling Cumulative Semantic Interference

Howard *et al.* (2006) presented an elegant model of the effect of cumulative semantic interference on response time. In their model, shared activation is implemented by assuming that words receive continuous (integrated over computationally discrete timesteps) activation from semantic nodes, and each time one semantic node is activated, similar semantic nodes are also activated to a lesser degree. Lexical competition is implemented by inhibitory connections running from each word to every other word (*i.e.* lateral inhibition). A word is only selected upon reaching an absolute selection threshold, but since activated words inhibit each other, strong competitors can slow down the selection of a target word. Finally, each time a word is selected, its connection from its semantic node grows stronger, implementing a priming function.

The Howard *et al.* model is noteworthy because it instantiates the principles of shared activation, competitive selection, and priming and because it attributes the interference to processes that are insensitive to time and to unrelated interference. Our goal is to extend this approach. We do so in three respects. First, we identify the priming mechanism with error-driven connectionist learning. This learning mechanism has the natural property that each act of retrieval in a certain context strengthens the target of retrieval (repetition priming) while at the same time making it less likely that similar memories are retrieved instead in that context (similarity-sensitive interference). We will show how this mechanism is consistent with both the perseveratory gradient in the production of word errors and the insensitivity of cumulative semantic interference to unrelated items or the passage of time. By attributing these effects to error-driven learning, we link up with the many cognitive models that are based on such learning (e. g. Chang *et al.*, 2006; Gupta & Cohen, 2002; Plaut, McClelland, Seidenberg, & Patterson, 1996) and, as we later demonstrate, this attribution addresses the question of whether retrieval-induced forgetting is caused by “inhibition”. Second, we develop the model in conjunction with theories of word production so that it can account for errors as well as response times. This requires a decision process that allows for lexical competition to play out in time, and for errors of commission and errors of omission. This proposed decision process may underlie the correlation between cumulative semantic interference and activation of Broca’s area. Finally and most importantly, we offer the hypothesis that cumulative semantic interference does not, in fact, require a competitive mechanism for lexical selection. Specifically, we demonstrate that competition in the lexical selection process is unnecessary when combined with error-driven learning. The resulting non-competitive model, we claim, can explain the major findings concerned with cumulative semantic interference.

### The Model

**Overview**—The key components of our model concern *lexical activation*, *lexical selection*, and *learning*. As in many models of lexical retrieval in production, retrieving a word begins with activating a set of semantic features (*e.g.* Dell *et al.*, 1997; Gordon & Dell, 2003; Rapp & Goldrick, 2000). These semantic features each connect to a number of words, and thus activate those words in proportion to the strength and number of these connections (lexical activation, Figure 1). Thus, multiple words are activated, requiring some kind of decision. For the model, we assume that the most active word is chosen. However, when more than one word is activated, it is assumed to be difficult to identify the most active one (*i.e.* if the difference in activations is slight, the winner is hard to “see”), so a ‘booster’ mechanism kicks in to tease the activations apart. This booster repeatedly amplifies each word’s activation until a winner can be selected (lexical selection), or until this boosting process times out. Response time is assumed to be correlated with the number of boosts needed for the winner to emerge. Errors of commission, such as semantic errors, occur when the wrong word is chosen, and errors of omission occur if the booster times out. Finally, after lexical

selection has concluded, an error-driven learning process adjusts the semantic-to-lexical connections so as to facilitate future retrieval of the target word (learning). In the following sections, we describe the details of model's architecture, and its lexical activation, lexical selection, and learning

**Model architecture**—The model is a feedforward two-layer network. Semantic feature nodes (*e.g.* FURRY or AQUATIC) form the input layer of the network. Each feature node connects directly to each of the word nodes (such as DOG or BOAT) in the output layer. Connection weights are initialized at zero and are continually adjusted through an error-driven learning process, as detailed later in this description. There are no lateral connections between semantic feature nodes or between word nodes, or reverse connections from words to features.

### Algorithms

**Lexical activation:** When semantic features are activated (as we assume happens when a picture is presented), these features in turn activate words. The net input,  $net_i$ , to any lexical node  $i$ , sums the activation,  $a_j$ , of each semantic feature,  $j$ , times the weight of its connection to the lexical node,  $w_{ij}$  (Equation 1).

$$net_i = \sum_j w_{ij} a_j \quad \text{Equation 1}$$

This net input,  $net_i$ , is then converted to an activation,  $a_i$ , via a logistic function (Equation 2).

$$a_i = \frac{1}{1 + e^{-net_i}} \quad \text{Equation 2}$$

Thus, the activations range from zero to one. We assume that lexical activation is imprecise and therefore add a small amount of normally-distributed noise,  $v$  (with a mean of 0 and a standard deviation of  $\theta$ ), to the net input,  $net_i$ , yielding Equation 3.

$$a_i = \frac{1}{1 + e^{-(net_i + v)}} \quad \text{Equation 3}$$

**Lexical selection:** The next stage applies a competitive winner-take-all process to the lexical activations, linking increased lexical competition to increased naming latencies. A booster mechanism floods the network with additional activation that combines nonlinearly with the existing lexical activation until either one word grows discernibly more active than the rest or the boosting process times out. Notice that this booster process is “dumb” in the sense that it does not know which word is the target. It repeatedly boosts all words. But because it boosts them in a multiplicative manner, the most active one gradually increases its lead on the other words.

The booster is engaged only to the extent necessary to select a single word (that is, it operates more when selection is difficult), recalling Schnur *et al.*'s (2009) reports of greater LIFG activity as a function of increased lexical competition. Therefore we tentatively identify this booster with the competition-biasing mechanisms that are hypothesized to be a function of the LIFG (*e.g.* Kan & Thompson-Schill, 2004; Thompson-Schill, D'Esposito,



Aguirre & Farah, 1997), but we acknowledge that our implemented booster has arbitrary properties that lack neural motivation. That is, we commit to the *functions* of the booster (aiding selection when competition is present) and its possible association to the LIFG (Broca's area), rather than to its implemented details.

The boosting process plays out over time. To determine whether a winner has emerged, at each timestep,  $t_n$ , we compare the difference between the activation of each word node,  $a_{i t_n}$ , and the mean activation of the other word nodes,  $a_{others t_n}$ , to a threshold value,  $\tau$  (Equation 4).

$$\tau > (a_{i t_n} - a_{others t_n}) \quad \text{Equation 4}$$

If no word's activation difference exceeds the difference threshold (*i.e.* Equation 4 is false for all  $i$ ), then the booster multiplies each word's current activation level,  $a_{i t_n}$ , by a constant boosting factor,  $\beta > 1.0$ . The result becomes its new activation level,  $a_{i t_{n+1}}$  (Equation 5). Then this testing and boosting process repeats.

$$a_{i t_{n+1}} = a_{i t_n} \beta \quad \text{Equation 5}$$

A word is selected, that is, the boosting stops, if and when its activation advantage over other words (per Equation 4) is great enough. The timestep at which this selection occurs,  $t_{selection}$ , is treated as an index of the duration of the lexical selection process, which should correlate with naming latency. If, for the sake of simplicity, we assume no variation in the repeated boosting, this iterative process becomes computationally equivalent to Equation 6.

$$t_{selection} = \log_{\beta} \left( \frac{\tau}{a_{i t_1} - a_{othres t_1}} \right) \quad \text{Equation 6}$$

However, if no node reaches the difference threshold within a certain number of boosts,  $\Omega$ , then no word is selected and the trial is an omission. This corresponds to a simple "wait and give up" theory of omissions. So we do not consider an omission as a special state that may be achieved, but rather a lack of sufficient evidence for any particular word, making it difficult to select a word quickly enough.

Note that while the implemented boosting process may be deterministic, based on the initial activations, the target word will not necessarily be selected. The combination of a discernible-difference threshold and a selection deadline may preclude selecting any word if the difference in lexical activations is too small. Adding noise to the lexical activations, as we have done, increases this chance and further opens the possibility that a competitor will be selected instead. Furthermore, although we have not done so here, one could assume that the boosting process is subject to noise either in its normal operation, or in pathological cases (*e.g.* LIFG damage), by allowing for boosts to randomly fail for particular words at particular time steps. A noisy booster would then have properties in common with sequential stochastic decision mechanisms such as a random-walk process.

We do not implement any residual activation or inhibition in this model. When the trial ends, either by selecting a word or by failing to select a word before the deadline, all activations return to zero.

**Learning:** At the end of each trial, semantic-to-lexical connection weights are adjusted according to Equation 7, which is the Widrow-Hoff or delta rule tailored for the logistic activation function (Widrow & Hoff, 1960; Rumelhart & McClelland, 1986):  $\Delta w_{ij}$  is the weight change for the connection to node  $i$  from node  $j$ ,  $\eta$  is the learning rate, and  $d_i$  is the desired activation of node  $i$ .

$$\Delta w_{ij} = \eta(a_i(1 - a_i)(d_i - a_i))a_j$$

Equation 7

Since this equation will prove crucial for understanding the behavior of the model, we should unpack it a bit more. We have said that learning is error-driven. This means that connections are adjusted according to  $(d_i - a_i)$ , the discrepancy between the desired activation of output node  $i$ ,  $d_i$ , and its actual activation,  $a_i$  (that is, its activation before boosting). So the error in a receiving node's activation affects both the degree and the direction of the weight change. Hence, when the error for an output node  $(d_i - a_i)$  is strongly positive, connections feeding it will be greatly strengthened. When the error for an output node is strongly negative, the connections feeding it will be greatly weakened. Notice that because the logistic activation function precludes activations that are actually 0 or 1, every word unit will experience at least some error on all trials, either positive or negative, if the desired activations are 0 or 1. Next, including the  $a_i(1 - a_i)$  component scales weight adjustments to  $a_i$ , such that weight changes are greatest at  $a_i = 0.5$ , and decrease as  $a_i$  approaches 0 or 1. Thus, weight changes are strongest for connections that contribute to moderate activations. Adding the  $a_j$  specifies that connections from that connections from input  $j$  should only be modified to the extent that  $j$  is activated. And finally, the learning rate,  $\eta$ , is simply an arbitrary global parameter, used to adjust how rapidly weight changes occur.

Thus the learning algorithm increases the connection weights from active semantic features to the target word, and decreases weights from those features to all other words, to the extent that those words were active before boosting. Since this learning is based on the deviation between  $d_i$  and  $a_i$ , it occurs regardless of whether the target was ultimately selected. So if the network encountered a dog (activating semantic features MAMMAL and TERRESTRIAL), then the connections from MAMMAL and TERRESTRIAL to DOG would strengthen, and the connections from MAMMAL and TERRESTRIAL to any other activated words (e.g. BAT) would weaken. The next time the network encounters a dog, those same semantic features will activate DOG more efficiently (i.e. activating DOG more and competitors less), increasing the speed and likelihood of its selection.

## Simulations

Our lexical learning model integrates several features common to theories of lexical access: lexical retrieval begins when distributed semantic features activate words (e.g. Dell *et al.*, 1997; Rapp & Goldrick, 2000); lexical selection uses a differential threshold (e.g. Levelt *et al.*, 1999); and semantic-to-lexical connections are adjusted through experience (e.g. Gordon & Dell, 2003; Howard *et al.*, 2006).

Can this model account for the major behavioral manifestations of cumulative semantic interference? To answer this question, we simulate several of the experiments described in the Introduction. First, Simulation 1 compares predicted selection latencies in a continuous paradigm simulation to Howard *et al.* (2006)'s behavioral data, and establishes one additional prediction from the model. In Simulation 2 we compare predicted latency patterns from blocked-cyclic presentation to data from Schnur *et al.* (2006, Experiment 1), Damian and Als (2005), and Belke (2008). Simulation 3 tests whether the simulated semantic blocking effect generalizes to new items, as reported by Belke *et al.* (2005). In Simulation 4,

we turn to the aphasic patient data. Repeating the procedure from Simulation 2 with noisier lexical activations, we compare the resulting error patterns with those reported by Schnur *et al.* (2006, Experiment 2) and Hsiao *et al.* (2009). Simulation 5 explores the mechanisms behind the model's effects by distinguishing the influence of weight increases (facilitatory learning) and weight decreases (inhibitory learning) on response times, analogous to the 'blocking' versus 'inhibition' debate in the retrieval-induced forgetting literature. Finally, Simulation 6 examines the role of competitive lexical selection in creating cumulative semantic interference, demonstrating that competitive lexical selection is not, in fact, necessary for any of the effects seen in Simulations 1–5.

Our goal in these simulations is to explore how a simple, omnipresent process – incremental learning – can lead to a surprising range of behavioral effects depending on the nature and ordering of stimuli during an experiment. Toward that end, we hold all aspects of the simulations constant throughout this paper, except where we have motivated reasons to change them.

We implement a standardized vocabulary structure for these simulations (Figure 2). Equal numbers of words share each semantic feature, and each word is uniquely specified by the conjunction of two semantic features. For example, WHALE, BOAT, and WATERLILY share the feature AQUATIC, whereas BAT, PLANE, and ORCHID share AERIAL. And the features AQUATIC and VEHICULAR specify the word BOAT, while AQUATIC and MAMMALIAN specify WHALE.

When simulating picture naming in the model, we assume that the target picture is correctly recognized and that its semantic features are properly activated. Features that should be active get activations of 1, and those that should not be active get activations of 0. So we do not simulate errors in pre-lexical processes, though we concede that they can occur (*e.g.* Rogers et al, 2004).

Each simulation consists of two phases: training and testing. In training, we simulate the acquisition of subjects' pre-experimental lexical-semantic knowledge. Then, in testing, we simulate the learning that occurs during the experiment. We present trials that mimic the experimental conditions being simulated and continue to adjust the connections according to the same learning algorithm (including parameter values) that were used during training. So the only difference between the training and testing phases is that the testing phases focus on particular subsets of the vocabulary.

When simulating multiple testing conditions, such as homogeneous versus mixed blocks, we want to compare these directly. So we start the testing phase for each condition with the same trained weights. Thus the only difference between the conditions is the semantic relationship of the items cued in their testing phases.

To simulate testing multiple subjects with pictures from many categories, we repeat each simulation 10,000 times. Individual differences in experience are simulated by beginning each replication with a fresh network in *tabula rasa* state, and then training it with 100 randomly ordered sweeps through the vocabulary to represent the experience that this model subject brings to the experiment. Variation in the networks' performance thus comes from activation noise (present in both training and testing) and differences in the order in which words are cued.

### Simulation 1 – Continuous Paradigm

Howard *et al.* (2006) reported that picture-naming latencies increased by a consistent amount for each same-category item that was named. The magnitude of the incremental

increase was unaffected by the number of intervening items from different categories. They claimed that shared activation, competitive selection, and priming were necessary for any model to account for these findings. Because our model implements all three of these properties, it should exhibit a lag-insensitive incremental increase in selection times.

In Simulation 1a, we approximate Howard *et al.*'s protocol by cueing for production of five items from a single semantic category, like "farm animals", one at a time. A variable number of unrelated fillers (two, four, six, or eight) are cued between each critical item and the next, allowing examination of how the accumulation of semantic interference is affected by the number of intervening filler items.

**Method**—Parameters for this simulation are given in Table 1. Constraints on these parameters are discussed in the Results section.

**Training:** Networks were trained, as described above, on a vocabulary of 20 semantic features and 50 words.

**Testing:** A list of 25 pictures comprised the test phase. Five of the items on this list were critical items and came from the same category. Each shared one of its two semantic features with every other critical item. Following Howard *et al.* (2006), a lag of 2, 4, 6, or 8 fillers separated each critical item and the next, making a total of 20 fillers. These fillers were randomly selected with the constraint that they shared no semantic features with the critical items; they were, however, free to share features with other fillers. Each lag occurred once in each list, yielding  $4! = 24$  possible lag sequences. Each of 10,000 networks was tested with each of these 24 sequences.

**Analyses:** Following Howard *et al.* (2006), mean lexical selection times were calculated for each lag and ordinal position. In our simulations, these selection times are the mean numbers of 'boosts' needed before one output node could be selected, as described in Equation 6. Following the standard practice in picture naming studies, only the selection times for correct responses were included in these calculations. Errors of omission and commission occurred in less than one percent of the trials and were excluded from the analyses.

**Results/Discussion**—As expected, each critical item took longer to select than the previous one, with each item contributing an equal increment to the selection time. There was no systematic variation in this effect over the different lags (Figure 3a). These results are entirely consistent with Howard *et al.*'s human data (Figure 3b).

The proper behavior of the model requires some reasonable limits on its parameters. For instance, the boosting factor ( $\beta$ ) must be greater than 1.00. Otherwise, the booster isn't a booster. The threshold ( $\tau$ ) and deadline ( $\Omega$ ) should be such that, given the value of the boosting factor, words are often selected before the deadline. Finally, the learning rate ( $\eta$ ) must be sufficiently large that its effect is not obscured by activation noise ( $\theta$ ).

The learning rate matters because incremental learning underlies the accumulating interference in this model. Each time a word is retrieved, connections from the activated semantic features to the activated words are adjusted. Connections supporting the target word strengthen and those supporting competitors weaken. This learning event promotes repetition priming during the subsequent retrieval of the same target from these semantic features. But if one of these features is instead used to cue a different target word, this same learning means the new target will be weaker and the previous target – now a competitor – will be stronger. So the new target is retrieved more slowly. When sequentially retrieving several words that all share a semantic feature, the learning that follows each retrieval event

makes the retrieved word a strong competitor and further weakens those competitors that have not yet been named. Each newly named word from the category therefore takes a bit longer to retrieve than the previous one. Thus the incremental increase in selection times arises from the incremental adjustments to the semantic-to-lexical connections.

Priming by error-based connection adjustments also makes the model's semantic interference effect insensitive to both time- and item-based lag manipulations. Time-lag insensitivity comes from the fact that all priming in the model is persistent, at least in the sense that the weight changes do not decay with time. Furthermore, there is no residual activation to decay while unrelated items are named. Robustness to long item-lags with intervening unrelated fillers comes from the learning algorithm. The adjustments that follow each trial only affect connections from the specific semantic features that were active during it. DOG, for instance, relies on connections from TERRESTRIAL and MAMMAL. The only thing that would then make it more difficult to retrieve DOG via these two features would be changes to connections from them. So as long as TERRESTRIAL and MAMMAL are not activated en route to lexical selection in any filler trials, no number of filler trials will ever affect DOG's intentional retrieval.

In the introduction, we asserted that incremental learning has both a dark side (cumulative semantic interference) and a light side (repetition priming). However, the first simulation only addressed the dark side. To see how these light and dark sides interact, we turn to a related experiment by Navarrete *et al.* (2008). They added a repetition component to the continuous paradigm. Participants named related pictures separated by fillers, but the third or fourth ordinal position was filled by either a novel related picture or the same picture that appeared in the first or second ordinal position, respectively. The question here was how this repetition would affect naming latencies for subsequent novel related items. Recall that Howard *et al.* (2006) showed that naming latencies increase linearly each time a related item is named. Now if cumulative semantic interference is *type-based*, meaning that what matters is the number of unique related pictures a person has named, then repeating one of the items should not increase RTs at all. However, if cumulative semantic interference is *token-based*, meaning that what matters is the number of times a person has named related pictures (unique or not), then there should be no difference between the interference that accrues from repeating an item and that which accrues from accessing another novel item. They found the latter: Each related retrieval, whether introducing a novel item or repeating an item previously accessed, contributed separately to the cumulative semantic interference.

But what of the model? Is its cumulative semantic interference effect type-based or token-based? To address this question, we now simulate Navarrete *et al.*'s (2008) repetition experiment as Simulation 1b.

**Method**—Parameters, vocabulary, and training for this simulation were identical to those of Simulation 1a. The testing phase differed slightly, following Navarrete *et al.* (2008).

**Testing:** The testing phase was identical to that of Simulation 1a, except in two respects. First, we included filler items between the critical items, but did not manipulate item-lag. Second, either the third or fourth ordinal position represented a second cueing of one of the earlier critical items. In one set of lists, the item in the first ordinal position was repeated in the third position. In the other set of lists, the item in the second position was repeated in the fourth position. So a five-position critical item sequence might go DOG BAT DOG WHALE MOLE, or DOG BAT WHALE BAT MOLE.

**Analyses:** Analyses followed the methods outlined for Simulation 1a. Instead, we compare the mean response times in sequences with a repeated item in ordinal position three to those with repetition in repeated item in ordinal position four.

**Results/Discussion—**Repeating one of the critical items produced additional interference, just like accessing another novel item from the same semantic category (Figure 4). Specifically, we see that the linear increase in response times as a function of ordinal position, apparent in Figure 3, grows normally when one of the ordinal positions is filled by a repeated item. This simulated finding mirrors Navarrete *et al.*'s empirical data. Thus, the model's interference effect, like that in the human data, appears to be token-based. In the model, this happens because error-based learning creates additional interference each time a related item is accessed.

It is also worth noting that the model approximates the relative sizes of benefit due to repeating an item and the cost that each similar item imparts, the former being several times larger than the latter. However, we do not want to emphasize exact quantitative properties of the model, as its representations and vocabulary are quite simplified.

### Simulation 2 – Blocked-cyclic Paradigm

In the continuous paradigm, each subject named a large number of pictures just once. The blocked-cyclic naming paradigm, in contrast, requires subjects to repeatedly name a small number of pictures. Here we gauge semantic interference effects by comparing blocks of pictures from the same semantic category, the homogeneous condition, to blocks of unrelated pictures, the mixed condition.

There are two major findings from this paradigm. First, when pictures are presented in the homogeneous condition, they take longer to name than the same pictures presented in the mixed condition (*e.g.* Damian *et al.*, 2001). Second, the magnitude of this semantic blocking effect increases with each cycle (Belke, 2008, Experiment 1; Belke *et al.*, 2005, Experiment 1; Damian & Als, 2005, Experiment 4; Schnur *et al.*, 2006, Experiment 1). So the blocked-cyclic paradigm's semantic interference effect grows with each cycle, similar to the way that the continuous paradigm's semantic interference effect grew with each related item. Our second simulation simulates this blocked cyclic procedure.

**Method—**Model parameters were identical to Simulation 1, and are given in Table 1.

**Vocabulary:** The vocabulary for this and the remaining simulations consisted of 12 semantic features mapped onto 36 words. Each feature cued exactly six words, and each word was cued by the intersection of exactly two features. Training followed the format described in the introduction to the simulations.

**Testing:** Two parallel testing phases simulated the homogeneous and mixed conditions of a blocked-cyclic naming experiment. In each condition, a set of six words was repeatedly cued for four cycles, for a total of 24 trials, with words ordered randomly within each cycle. The homogeneous condition used six words from a single category, so they all shared one of their two semantic features. Words in the mixed condition each represented a different category, so none shared a feature with any other word in the set.

Both testing phases began with the same trained connection weights. But learning then continued separately in each condition. This learning proceeded at the same rate ( $\eta$ ) as during the shared training phase. So each replication tests the same trained network in both conditions, with no carryover from one condition to the next.

**Analyses:** We report selection times in terms of the mean number of ‘boosts’ for each item position in each condition. As before, errors of omission and commission occurred in less than one percent of the trials and were excluded from the analyses.<sup>2</sup>

**Results/Discussion—**The trial-by-trial predictions (Figure 5a) show both repetition priming and cumulative semantic interference. Repetition priming occurs in both the homogeneous and mixed conditions; each time a word is cued, it is selected more quickly. This is because the incremental learning that follows each retrieval facilitates future retrievals of the target word by strengthening its connections and weakening connections from its features to competitor words. Cumulative semantic interference only occurs in the homogeneous condition, creating an incremental increase in selection times within each cycle. This is the same pattern that we saw in Simulation 1, because it arises from the same process. However, in this simulation, the relation between repetition priming and cumulative semantic interference becomes more apparent. Each time one word gets stronger, its competitors get weaker. When that word is retrieved again, it is primed. When one of the competitors is retrieved, though, it is subject to interference. Because repetition priming and cumulative semantic interference are both at work in the blocked-cyclic paradigm we get a sawtoothed function for selection times in the homogeneous condition (Figure 5a), and a small decrease in the per-cycle mean selection times (Figure 5b). A smaller per-cycle decrease is precisely what we see in human data for the same task (*e.g.* Figure 5c, from Schnur *et al.*, 2006, Experiment 1). So it appears that the model successfully extends to blocked-cyclic naming.

Our model predicts some attenuation of the repetition priming across cycles, because of the learning algorithm. Each retrieval is a learning event, thus decreasing error for the next retrieval. And since the magnitude of each weight change depends on the magnitude of the erroneous activation for the relevant output node, this error-reduction reduces the weight change (and hence repetition priming) that will result from the next repetition. We acknowledge, though, that with our current parameter values for the learning and decision processes, the non-linearity is smaller than it is in the human data.

There is one feature of the data that this simulation does not exhibit. While our main effect of blocking begins in the first cycle, Belke *et al.* (2005) observed that in human data it tends not to appear until the second cycle. In fact, human data sometimes shows a brief semantic facilitation effect, for example in the first cycle depicted in Figure 5c. However, there is evidence that this early facilitation represents a conscious strategic process rather than an integral part of lexical retrieval, and is therefore beyond the scope of a model of lexical access. Three findings support this conclusion. First, Wheeldon and Monsell (1994) reported a similar brief semantic facilitation in a continuous paradigm experiment, and argued that it followed a different time course than the longer-lasting interference effect, suggesting a separate process. Extrapolating to the blocked-cyclic paradigm, a brief semantic facilitation could delay the appearance of an interference effect until the second cycle. Second, Damian and Als (2005, Experiment 4) demonstrated that the semantic interference appears in the first cycle, and grows thereafter, if homogeneous and mixed sets are interwoven a single block. This suggests that subjects’ expectations may play a role in the early facilitation. Finally, and most conclusively, Belke (2008, Experiment 1) showed that adding a working memory load to the blocked-cyclic naming paradigm led to semantic interference in the first cycle, with no sign of early facilitation. Since the dual-task disrupts semantic facilitation while leaving the growing semantic interference effect intact, we can conclude that the

---

<sup>2</sup>Errors were rare, but we note that they were relatively more frequent in the homogeneous condition than the mixed, and in both conditions, the errors were predominantly omissions. Given the infrequency of errors in this simulation, however, we do not discuss them further here.

facilitation represents a resource-demanding process that is distinct from whatever underlies the cumulative semantic interference effect and is not integral to the process of lexical retrieval.

### Simulation 3 – Generalization of Interference to New Pictures

An interesting empirical property of the semantic blocking effect is that it extends to naming new pictures from the same category (Belke *et al.*, 2005, Experiment 3). For instance, repeatedly naming a small set of birds, such as CROW, FINCH, and GULL, creates a substantial semantic blocking effect that will carry over, without interruption, to naming an entirely new set of birds. Belke *et al.* (2005) argued that this carryover reflected a refractory behavior where related words become temporarily inaccessible due to residual activation and/or inhibition.

In order to test whether a learning model could account for Belke *et al.*'s finding without representing residual activation or inhibition, we now test the model using their procedure. They compiled sets of pictures of from one category, and had subjects name half of these pictures for four cycles before switching over to name the other half for four cycles. For example, given a set of birds {CROW FINCH GULL JAY ROBIN SPARROW} subjects might name CROW, FINCH, GULL, FINCH, CROW, GULL, FINCH, GULL, CROW, GULL, FINCH, CROW, and then immediately switch over to JAY, ROBIN, SPARROW, ROBIN, SPARROW, JAY, ROBIN, JAY, SPARROW, JAY, SPARROW, ROBIN. The crucial question is whether the naming latency difference between related and unrelated conditions that develops while cycling through the first subset will continue in the first cycle of the new subset.

**Methods**—We followed the methods of Simulation 2, with just one change in the testing procedure. Instead of cuing one set of six words for four cycles during testing, we now cued one set of three words for four cycles, and then a set of three different words for four additional cycles. In the homogeneous condition, all six words shared a single semantic feature. None of the six words in the mixed condition shared any semantic features.

Model parameters were identical to Simulation 1 and 2, and are given in Table 1.

**Results**—The simulated blocking effect carried over from the first to the second subset without interruption (Figure 6a). Cumulative semantic interference built up while naming the first subset in the first four cycles, resembling the findings from Simulation 2. Upon switching to the second subset, in cycles five through eight, the interference effect continued with the entirely new items. The blocking effect in the fifth cycle exceeded that in the first cycle, demonstrating that the accumulated semantic interference transferred to new items from the same category.

In the model, interference generalizes to new items from the same category because the incremental weight changes that follow each selection affect all competitors. Indeed, this is required for the model to simulate the results of Howard *et al.*'s (2006) continuous paradigm in which same-category items do not repeat. Importantly, the fact that our simulated data (Figure 6b) mirrors Belke *et al.*'s (2002) human data (Figure 6c) demonstrates that the human results do not depend on the residual activation or inhibition that are normally associated with refractory behaviors (*e.g.* Forde & Humphreys, 1997; McCarthy & Kartsounis, 2000). Rather, this behavior can arise from persistent incremental learning.



### Simulation 4 – Aphasic Errors

Semantic blocking manipulations elicit longer picture naming latencies from healthy subjects, and they also affect lexical selection errors made by individuals with aphasia (Schnur *et al.*, 2006). Semantic errors and omissions become increasingly likely in the homogeneous condition, compared to the mixed baseline (Schnur *et al.*, 2006, Experiment 2). Other types of errors (*e.g.* phonological or unrelated errors) do not show such effects. Furthermore, when patients name pictures incorrectly, their within-set substitutions tend to match the words that they have used more recently, and there was no difference between this perseverative recency effect at one-second and five-second inter-stimulus intervals (Hsiao *et al.*, 2009).

To simulate the patient error effects, we need a theory of the differences between aphasic and nonaphasic lexical access in this task. Here we follow the approach of several researchers (*e.g.* Dell *et al.*, 1997; Rapp & Goldrick, 2000) and attribute impaired performance to the alteration of one or more model parameters, but without changing any of the model's processes. Specifically, we increase the activation noise parameter, with the result that errors become much more likely. Rapp and Goldrick specifically simulated aphasia by activation-noise lesions, and Dell *et al.* demonstrated that noise lesions mimic the decay lesions that they used to simulate some patients.

Schnur *et al.* (2006) compared unimpaired (Experiment 1) and aphasic (Experiment 2) performance in the blocked-cyclic paradigm by running both groups in essentially the same experiment. We do likewise, repeating Simulation 2, which we had based on their Experiment 1 procedure, with noisier lexical activations. Our analyses follow Schnur *et al.* (2006, Experiment 2) and Hsiao *et al.* (2009), so that we can compare the model's predictions directly to the data. So, if successful, the predictions should show Schnur *et al.*'s semantic blocking effects for semantic errors and omissions, and Hsiao *et al.*'s recency gradient for perseveration errors.

**Methods**—The vocabulary, training, and testing followed Simulation 2, only increasing the amount of noise in the lexical activations from 0.5 to 1.0. The new parameters are given in Table 2.

The analyses followed those that Schnur *et al.* (2006) and Hsiao *et al.* (2009) used for their human data. Schnur *et al.*'s (2006, Experiment 2) major findings concerned omissions and semantic errors, so our analyses focused on these as well. As described in the Model Description section, omissions occurred when the activations of lexical nodes were so similar that no winner could be selected before the deadline. Incorrect selections were classified as either *semantic errors* or *other errors*, according to whether the target and error shared a semantic feature. Following Schnur *et al.*, we calculate per-cycle means for each error type.

Perseveration error analyses followed Hsiao *et al.* (2009). These errors occurred when an erroneously selected word matched another target in the block whose name had been selected previously. These errors were rare in the mixed condition, and so perseveration error analyses were restricted to the homogeneous condition. For each such error, we recorded the number of trials back that the erroneously selected word was most recently selected. So if a dog was named as DOG, a bat was named as BAT, and a whale was then named as DOG, then the whale → DOG error would be counted as a perseveration at lag-2 because DOG was last produced two items back. Actual target-error pairs from the testing phase were then randomly reshuffled within each cycle, with the results coded as above, in order to estimate the probability that such a lag distribution might occur by chance.

Following Hsiao *et al.*, (2009, pp.136–137), and Cohen & Dahan (1998, p. 1643) before them, we thus derived a chance-corrected estimates of perseveration frequencies at each lag.

Hsiao *et al.* (2009) reported perseveration functions at both two ISIs. Our model assumes no time-based decay, and should therefore predict the same effects for any ISI. However, we give some measure of the variability of our predicted perseveratory effect by presenting data from four 10,000-block replications.

**Results**—Doubling the activation noise substantially increased the predicted error rates, producing errors in 12.4 percent of homogeneous trials and 10.9 percent of mixed trials.<sup>3</sup> Errors of commission were largely semantic, with less than 0.01 percent unrelated errors occurring in either condition.<sup>4</sup> The homogeneous condition elicited more semantic errors (Figure 7a) and omissions (Figure 7b) than the mixed condition. These effects increased across cycles, recalling Schnur *et al.*'s (2006, Experiment 2) aphasic patients' error patterns (Figure 7c).

Within-set substitution errors accounted for 51 percent of the semantic errors in the homogeneous condition, in agreement with Schnur *et al.* who reported 53 percent. In our simulated data, most (82.2 percent) of these within-set substitutions were identified as perseverations. As illustrated in Figure 8a, the tendency to perseverate a word exceeds chance at short lags, but falls to near-chance levels thereafter. This pattern recalls Hsiao *et al.*'s (2009) report that aphasic patients' were more likely to perseverate words that they had used more recently (Figure 8b).

One apparent difference between our predicted lag function (Figure 8a) and Hsiao *et al.*'s (2009) mean lag function (Figure 8b) is that our model is most likely to perseverate words at lag-1, whereas their patients were not. However, Hsiao *et al.* noted that this was not actually a reliable feature of their patient data, and actually excluded lag-1 from their statistical analyses. Roughly half of their patients showed this lag-1 dip, while half did not. Given this lack of consistency across patients, we follow one of Hsiao *et al.*'s proposed explanations, positing that their lag-1 dip may merely reflect a repetition-avoidance strategy that some patients chose to employ. Such a strategy has some precedence in non-patient work. For instance, Anderson and Neely (1996) cite two early naming-to-definition experiments demonstrating that preceding a trial with a semantically related prime produced facilitation if the prime was never the correct answer, but interference if the prime was sometimes the correct answer (Brown, 1979; Roediger, Neely, & Blaxton, 1983). This change suggests that a brief facilitation or anti-perseveration effect may merely index participants' sensitivity to the sequential-statistical properties of the testing paradigm (*i.e.* participants learn to expect that no picture will be immediately repeated, allowing them to constrain their predictions for the next item). Consistent with applying such an interpretation to this patient data, Hsiao *et al.* noted that targets only immediately repeated in 2.2 percent of their trials, so a repetition-avoidance strategy would have helped performance on the vast majority of trials.

While the precise omission-to-commission error ratio depends on the selection deadline, all of the error predictions that we have presented here hold true for a wide range of parameters.

<sup>3</sup>These totals included 2.7 percent semantic errors and 8.9 percent omissions overall, but these relative proportions depend on the values of the selection deadline and activation noise parameters. That is, a greater value for the deadline would decrease the number of omissions and increase the number of semantic errors, so the fact that there are more omissions than semantic errors is merely an accident of parameters. Therefore we focus on these error types individually, but avoid comparing them.

<sup>4</sup>Schnur *et al.* (2006) also reported that "Other errors" (*e.g.* phonological slips and unrelated word errors) become increasingly likely in the Mixed condition, relative to their frequency in the Homogeneous condition, as shown in Figure 7c. Although our model does not currently generate such errors, adding some noise to semantic-input activations increases the likelihood of unrelated word errors, and these do in fact show such reverse blocking effects.

As long as there is sufficient noise to produce errors, but not so much that selection becomes completely random, the semantic error and perseveration effects emerge. And as long as the selection deadline parameter cuts off some, but not all, lexical selections, the omission effects also show up. Thus, the effects are not just a matter of fitting the model to the data.

So where do these error effects come from? Since activations return to baseline at the end of each trial, any errors necessarily result from the interaction of activation noise and persistent changes in connection weights. Recall that omissions occur when target and competitor activations are too similar for a winner to be resolved before the selection process times out. Semantic errors happen when a non-target word from the target category becomes substantially more active than its competitors, including the target. The conditions underlying these errors are rare, under normal circumstances. But adding noise to the model increases the likelihood that errors will arise from target and competitor net inputs that are merely “somewhat similar”. And these somewhat similar net inputs drove the selection latency effects in Simulations 1–3. Thus, the error effects derive from the same basic process that led to latency effects in the previous simulations: incremental learning.

Incremental learning (Equation 7) also drives the recency effect for perseverations. Here it may be helpful to work through an example. Let us imagine a patient sequentially naming three pictures, DOG, BAT, and WHALE, that all share a MAMMAL feature. When he first names DOG, the MAMMAL feature will support each of these words more-or-less equally. After naming DOG, however, the link from MAMMAL to DOG is strengthened and the links from MAMMAL to BAT and to WHALE are weakened. Thus, when the patient encounters the picture BAT, the MAMMAL feature will activate DOG more strongly than BAT or WHALE, making him more likely to name the picture as DOG than as WHALE. This is the primary basis of the recency effect: words tend to form stronger competitors when their connections from shared features have been strengthened more recently, and weakened less recently.

We should note, however, that other learning models could explain a recency effect, provided that they can explain errors in addition to response times. In Howard *et al.*'s (2006) priming model, for instance, a semantic-to-lexical connection is strengthened by a fixed increment each time its associated word is selected. Thus, the DOG-BAT-WHALE explanation for recency outlined above is also consistent with models that learn just by strengthening each item by a fixed amount.

Does our error-based learning algorithm offer any advantage in explaining the recency gradient for perseverations? To investigate this question, we repeat the perseveration simulation with three different learning rules applied at testing. First, we present results from a model using our standard delta rule learning algorithm. To gauge the contribution of error-based learning over other approaches, we compare these results to those from a version of the model that uses a non-error-based connection-strengthening algorithm. Finally, to assess how much of the perseveration effect might reflect random variation, structural peculiarities of the blocked-cyclic paradigm, or persistent biases that cannot be attributed to either learning process, we present results from a version that does not learn at all during the testing phase.

**Methods**—Networks were trained as normal, using a delta-rule learning algorithm and the parameters listed in Table 2. Perseveration simulations followed, as described above, but with two changes.

First, after training each network as normal, we applied one of three learning algorithms at the time of testing. The *delta-rule simulation* implemented the same error- and activation-

proportional learning algorithm (Equation 7) that we have used elsewhere in this paper. The *priming simulation* instead implemented a fixed-increment unsupervised learning algorithm. Each time a word was selected, connections to it from the active features were strengthened by a fixed increment.<sup>5</sup> Finally, the *non-learning simulation* did not strengthen or weaken any connections during the test phase.

Second, to preclude the possibility that omission errors might obscure differences in the perseveration effects, we removed the timeout parameter ( $\Omega$ ) during the test phase. So these simulations do not allow for omission errors, meaning that they will not be directly comparable to our previous simulations. However, removing the omission behavior allows us to keep all parameters constant across simulations, without further concern for matching error and omission rates.

**Results**—The delta rule and priming simulations produced equivalent rates of semantic errors (5.79% and 5.80%, respectively). This means that we can compare the strength of these recency effects more or less directly. The non-learning simulation produced slightly more semantic errors (6.60%), owing to the fact that its mappings never improved after training. As in the previous simulation, non-semantic errors were exceedingly rare, so we describe perseveration effects only for the homogeneous condition.

In the delta-rule simulation, perseverations were again much more likely to match recent responses (Figure 9). The priming simulation also produced some recency effect for perseverations, though this was much smaller. And the non-learning simulation demonstrates that both effects exceed what we might expect to emerge from any persistent biases in model weights. It is thus clear that incremental learning is crucial for the recency effect.

The delta rule creates the stronger recency gradient by causing connection weakening as well as connection strengthening. As in the priming account, the more recent a potential competitor is, the more likely that it will have been strengthened more times by the delta rule in the blocked-cyclic paradigm. However, the delta rule amplifies the gradient because the more recent a competitor is, the fewer times it will have been weakened by the action of the rule, since its last strengthening. So, in the sequence BAT, DOG, WHALE, when WHALE is the target, both BAT and DOG have both been strengthened, and so both are potential perseverates. But DOG, the more recent item, is a more likely error for WHALE because BAT was weakened during the DOG trial. BAT was especially active at that time (because it had just been strengthened), and hence the delta rule will have weakened it in proportion to its activation. Thus, DOG will be a more powerful competitor than BAT during the WHALE trial.

### Simulation 5 – Facilitation versus inhibition

Our learning algorithm specifies changes in connection weights whenever there is some degree of error and it does not distinguish between the learning that strengthens connections and the learning that weakens them. For example, we saw that, in the analysis of the model's perseverations, both strengthening and weakening of connections contribute to the effect of recency on these kinds of errors. But both processes arise from the same weight-change equation.

Several explanations of cumulative semantic interference, however, attribute it either to a process that strengthens competitors or a process that inhibits targets, but not both, and thus

<sup>5</sup>We set this increment equal to the average weight increase for the first target in the test phase of the delta rule simulation (*i.e.*  $\Delta w_{ij} = \eta(a_i(1 - a_i)(d_i - a_i))a_j = 0.75 \times (0.72 \times (1 - 0.72)(1 - 0.72)) \times 1 = 0.042$ ).

there is a debate about whether this effect is, at core, one of facilitation or inhibition. This is analogous to the occlusion versus inhibition debate in the retrieval-induced forgetting literature. The central idea in the facilitation (or occlusion) explanation is that retrieving an item facilitates its future retrieval, making it a stronger competitor when related words are cued. Thus, lexical competition grows more intense when retrieving words in a homogeneous context because each new target faces a growing number of pre-potent competitors. This facilitatory account dominates much of the cumulative semantic interference literature. Damian *et al.* (2001) and Belke *et al.* (2005) both suggested temporary facilitation mechanisms, while Wheeldon and Monsell (1994), Damian and Als (2005), Schnur *et al.* (2006), and Howard *et al.* (2006) argued for more persistent facilitation. Some memory researchers have also offered facilitation-based explanations of retrieval-induced forgetting (*e.g.* Macleod *et al.*, 2003; McGeoch, 1942; Mensink & Raaijmakers, 1988).

However, inhibition has played a more prominent role in discussions of retrieval-induced forgetting (*e.g.* Anderson, 2003; Anderson *et al.*, 1994; Postman *et al.*, 1968). The basic idea here is that retrieving a target necessarily harms its competitors, whether by inhibiting concepts directly (*e.g.* Postman *et al.*, 1968) or merely making them less accessible via retrieval cues (*e.g.* Melton & Irwin, 1940). For cumulative semantic interference, an inhibition account would entail the active suppression of lexical competitors during each naming attempt, with the lasting consequence of making them less accessible. Each new target in a homogenous set then becomes increasingly difficult to retrieve by virtue of being repeatedly suppressed each time one of its competitors is retrieved instead. Though pure inhibitory explanations are relatively rare in discussions of cumulative semantic interference, we note one example in McCarthy and Kartsounis' (2000) report of patients' omission errors.

So in all these cases we see semantic interference attributed to priming that either strengthens recent targets or weakens recent competitors, but not both. And, as Schnur *et al.* (2006) pointed out, these two mechanisms generate different predictions. We can use our model to separate the strengthening and weakening processes to see what each contributes to predicted cumulative semantic interference effects. In this simulation, we selectively apply either connection strengthening or connection weakening while disrupting the other. If the model realizes the notion that cumulative semantic interference arises from strengthening targets, then a version of the model that only allows for connection strengthening should exhibit the interference effect. If cumulative semantic interference arises from weakening competitors, then it should be seen in a model version that only allows for connection weakening.

**Methods**—Model parameters were identical to Simulation 1, 2 and 3, and are given in Table 1.

Methods for these simulations followed Simulation 2 exactly, but with one additional manipulation during the testing phase. Each replication began with the standard training phase, where learning was applied according to the delta rule. So the training was identical to Simulation 2. During the testing phase, however, the learning algorithm was modified to only apply either weight increases or decreases. In Simulation 5a, these weight changes only increased connection weights (Equation 8) and in Simulation 5b, they only decreased the weights (Equation 9).

$$\Delta w_{ij} = \begin{cases} \Delta w_{ij} & \text{if } \Delta w_{ij} \geq 0 \\ 0 & \text{if } \Delta w_{ij} < 0 \end{cases} \quad \text{Equation 8}$$

$$\Delta w_{ij} = \begin{cases} \Delta w_{ij} & \text{if } \Delta w_{ij} \leq 0 \\ 0 & \text{if } \Delta w_{ij} > 0 \end{cases} \quad \text{Equation 9}$$

**Results**—The simulations that only increased or only decreased weights had distinct effects on simulated lexical selection times (Figure 10). Recall that, in Simulation 2, selection times in the homogeneous condition increased within a cycle and decreased across cycles, while the blocking effect (homogeneous-mixed difference) increased both within and across cycles (Figure 10a). This pattern represents the combined effects of increased and decreased weights. In Simulation 5a, weight increases alone produced a step-like decrease for selection latencies in both homogeneous and mixed conditions (Figure 10b). This is what we would expect from a repetition priming effect. However, as shown by the close overlap of curves, the semantic blocking effect was minimal, suggesting that weight increases play a much stronger role in repetition priming than cumulative semantic interference. Learning through weight decreases alone (Simulation 5b) generated a robust semantic blocking effect with only a very weak repetition priming effect (Figure 10c). Selection times steadily increased in the homogeneous condition and decreased only slightly in the mixed condition.

Errors were rare, occurring in less than one percent of the trials in either simulation. But, as in Simulation 2, these errors were predominantly omissions and were relatively more likely in the homogeneous conditions than in the mixed. As one would expect from Figure 10, this blocking effect was stronger when learning via weight decreases than when learning via increases. Moreover, learning via weight decreases produced a blocking effect for omission errors that increased across cycles, thus tracking the interference effects in response times seen in Figure 10c.

From these findings, we can characterize the model's repetition priming as a facilitatory (strengthening of weights) effect, and its cumulative semantic interference as an inhibitory (weakening of weights) effect. Thus, the model's account clearly differs from a popular characterization of cumulative semantic interference as a facilitation-based effect, adopting instead the inhibition explanation more common in discussions of retrieval-induced forgetting.

### Simulation 6 – Competition in selection and learning

In the previous simulation we explored the contributions of weight strengthening and weakening to repetition priming and cumulative semantic interference. Crucially, we found that the weight-strengthening version of the model failed to produce cumulative semantic interference. This finding is unexpected because this restricted version still implements Howard *et al.*'s (2006) necessary and sufficient principles of shared activation, priming, and competitive selection. Shared activation was clearly present in the distributed semantic features. The experience-based weight increases facilitated access to recent targets, fulfilling Howard *et al.*'s priming function. And the lexical selection algorithm was competitive by virtue of implementing a differential threshold. So why didn't it work?

The answer may lie in the notion of competitive selection that our model implements. In a nutshell, our version of competitive selection may not have been competitive enough. Recall that our selection rule (Equation 6) compares each word's activation to the mean activation

of all its competitors. Facilitation-based explanations of cumulative semantic interference often focus on the role of individual prepotent competitors in prolonging lexical selection times. But, with a large vocabulary that includes lots of unrelated words, using an average activation for the differential threshold could minimize the impact of these prepotent competitors. That is, having a competition between the target and the mean activation of other words is quite close to having a competition between the target and some absolute standard or threshold, which is not technically a competition.

If this analysis is correct and lexical selection in our model truly functioned as a non-competitive process, then by Howard *et al.*'s logic the full version of the model with both strengthening and weakening of weights should not have been able to generate cumulative semantic interference in any of the previous five simulations. But it did. In the following simulations we explore this puzzle and offer a solution, a solution that forces us to change our conceptions of what constitutes competition. Specifically, we first verify our intuition that the competition involving the mean of the competitors is functionally like that of using a non-competitive absolute-threshold decision rule by repeating Simulation 5 with an absolute threshold. From this, we then develop a new hypothesis about the nature of competition and test one of its ramifications by repeating Simulation 5 with a clearly competitive selection rule in which the target is compared against only its most potent competitor.

**Methods**—We repeated Simulation 5 with an absolute (*i.e.* non-differential) threshold, replacing Equation 6 (reprinted below as Equation 8) with Equation 9, below. Thus, lexical selection becomes a non-competitive horse race. Competitors' activations do not affect selection times, and the first word to reach the threshold wins.

$$t_{selection} = \log_{\beta} \left( \frac{\tau}{a_{i t_1} - a_{othres t_1}} \right) \quad \text{Equation 8}$$

$$t_{selection} = \log_{\beta} \left( \frac{\tau}{a_{i t_1}} \right) \quad \text{Equation 9}$$

Otherwise, this simulation replicates Simulation 5 exactly.

**Results**—The results from the absolute threshold model (Figure 11) were virtually indistinguishable from those from the differential-threshold model that we used in the first five simulations (Figure 10). The full-learning condition produced the familiar saw-toothed function (Figure 11a), indicating repetition priming and cumulative semantic interference. Weight increases carried a repetition priming effect (Figure 11b), while decreases carried the interference effect (Figure 11c).

In fact, the absolute threshold version of the model is not just indistinguishable from the differential-threshold model with regard to the saw-tooth function. It can simulate every one of the phenomena that the differential-threshold model did. Figure 12 shows the results of new simulations that repeat simulations 1–4 while replacing the competitive selection rule with a non-competitive absolute threshold.

These results are important for two reasons. First, they fit our previous results to a tee. This consistency means that not only was competitive selection unimportant for our previous findings, but it did not even contribute to them. Second, demonstrating that we can get

cumulative semantic interference effects without a competitive selection mechanism proves that cumulative semantic interference does not require competitive lexical selection.

So how does the weight-reduction version of the model (Figure 8c) produce cumulative semantic interference without competitive selection? We claim that the interference effect requires competition in a broader sense, not necessarily limited to the selection process. If the essence of competitive selection is to cause the facilitation of one word to slow down retrieval of other words, then our model *already does that through its learning algorithm*. The learning process naturally involves weakening connections to competitors while strengthening connections to targets. Weakening connections slows down the subsequent retrieval of competitor, thereby creating cumulative semantic interference. So, in this context, error-based learning is competitive.

Howard *et al.*'s (2006) model required competitive lexical selection because it offered a wholly facilitation-based account of cumulative semantic interference. Their priming mechanism strengthened connections to targets without weakening connections to competitors. So their model implemented something like our facilitatory learning condition. And within those confines, they needed competitive selection to convert repetition priming into a competitor-inhibition effect and achieve semantic interference (*e.g.* Mensink & Raaijmakers, 1988).

We can test this analysis. Having established our original mean-based selection rule as essentially non-competitive, we can try making it more competitive. Instead of comparing each word's activation to the mean activation all its competitors, we might hone in on just the strongest. The idea here is to amplify the effect of the strongest prepotent competitor, so that it has a greater impact on selection times. Thus, we replace Equation 6 with Equation 10, below.

$$t_{selection} = \log_{\beta} \left( \frac{\tau}{a_{i t_1} - a_{strongest\ competitor\ t_1}} \right) \quad \text{Equation 10}$$

Now we repeat the previous batch of simulations. If this tweak of the competitive selection rule is sufficient to turn repetition priming into cumulative semantic interference, then we should see it in the facilitatory learning condition.

And we do (Figure 13). In the facilitatory learning condition, focusing competitive selection on a single close competitor translates the repetition priming into cumulative semantic interference (Figure 13b). Interestingly, the selection time pattern in this facilitatory learning condition resembles that of the simulation including both weight increases and decreases (Figure 13a). Thus, competitive selection allows facilitatory changes (as in Howard *et al.*'s model) to account for both repetition priming and cumulative semantic interference effects. We should note, however, that inhibitory learning (Figure 13c) still plays a major role in creating the combined interference effect, demonstrating its continued relevance to explanations of cumulative semantic interference.

The results of Simulations 5 and 6 suggest that cumulative semantic interference requires competition, but not necessarily during lexical selection. Incremental learning can also be competitive, and competitive *learning* is sufficient to produce cumulative semantic interference without competitive *selection*. Therefore, the phenomenon of cumulative semantic interference cannot be claimed to uniquely support the existence of a competitive mechanism for lexical selection in speech production. We expand on this important conclusion in the general discussion.



## General Discussion

### Summary and Implications of Findings

Lexical retrieval leads to lexical learning. The light side of learning is well known. Retrieving the same word again becomes faster and more accurate. But learning also has a dark, competitive, side that hinders the subsequent retrieval of semantically related words. In our theoretical framework, this dark side of learning leads to the behaviors identified with cumulative semantic interference.

Remarkably, this framework does not require competitive lexical selection. Competitive learning obviates the need for competition in the selection process. Thus, we can modify Howard *et al.*'s (2006) explanation to identify four properties that must be true of lexical retrieval in order to account for cumulative semantic interference: shared activation, activation-dependent selection time, persistent priming, and competition. First, activating one word must also activate similar words. Though, as Howard *et al.* (2006) pointed out, this *shared activation* can be accomplished in many ways, it is an inherent property of models such as ours that rely on distributed semantic representations. Second, more activated words should be selected more quickly than less activated words. Such *activation-dependent selection time* is common to most theories of lexical selection, both competitive (*e.g.* Roelofs, 1992) and non-competitive (*e.g.* Mahon *et al.*, 2007). Third, retrieving a word must persistently prime its future retrieval. Such *persistent priming* is most readily understood as incremental learning (*e.g.* Damian & Als, 2005). Finally, facilitating retrieval of one word must have negative consequences for related words. This *competition* could be implemented in many ways, including competitive lexical selection, but we note that it follows naturally from implementing priming as error-based learning.

Given these principles, our learning model offers a parsimonious account of the empirical aspects of cumulative semantic interference. It incorporates distributed semantic representations, error-based learning, and a booster mechanism that produces activation-dependent lexical selection times and errors.

### Response time effects

The model's selection times reflect all the response-time hallmarks of cumulative semantic interference. Lag-invariant incremental increases in response times, in the continuous paradigm (Simulation 1a), showed that the model can account for Brown's (1981) and Howard *et al.*'s (2006) main findings. Moreover, the model simulated the dual effects of repeating an item (Navarrete *et al.*, 2008): that item is named much more quickly (repetition priming), but repeating it creates additional semantic interference, just about as much interference as two different related items would (Simulation 1b). Applied to blocked-cyclic naming (Simulation 2), the model's selection times showed both repetition priming and cumulative semantic interference effects (*e.g.* Damian *et al.*, 2001; Schnur *et al.*, 2006, Experiment 1). And the accumulated interference transferred to novel items from the same semantic category (Simulation 3), recalling Belke *et al.*'s (2005) report. Thus, we have demonstrated that the model accounts for the major reaction time manifestations of cumulative semantic interference.

### Error effects

The learning model can also account for aphasic patients' error patterns (Simulation 4). One way of understanding aphasic brain damage is to assume that processes work normally, but are more prone to error. The model was made more error-prone by adding a small amount of noise to each word's net input. This noise led to blocking effects for semantic errors and omissions that increased with each cycle, reminiscent of findings from Schnur *et al.*'s (2006,

Experiment 2) patient work. And the model's perseveration errors recalled Hsiao *et al.*'s (2009) lag-based recency gradient, suggesting a further match to the empirical patient data.

Although we are heartened by the degree to which the model simulates the major response-time and error data effects in the literature, we acknowledge the model's limitations. We do not, indeed we can not, 'fit' the data quantitatively. This is because the model's lexical base is small and its treatment of semantics is rudimentary. Most humans, for example, know more than thirty-six words, and would describe a dog as something more than merely a terrestrial mammal. Moreover, we have constrained the scope of the model to lexical activation and selection. We have not represented any processes leading up to the activation of semantic features, nor any processes that follow lexical selection. And, although we postulate that the experimental paradigms that are modeled may induce strategies, we do not simulate these. All of these factors compromise the model's ability to fit the actual numbers. What is not compromised by the model's simplifications, we would argue, is its ability to make theoretical issues more transparent. This was a strength of Howard *et al.*'s (2006) model and we hope to achieve the same with our model, particularly with regard to the issue of the role of competition, to which we turn next.

### Competition and the occlusion versus inhibition debate

We provided an analysis of the model's account of cumulative semantic interference in Simulations 5 and 6. In contrast to several previous accounts, facilitatory processes contributed almost nothing to the model's interference effect (Simulation 5). Learning-based weakening of semantic-to-lexical connections, however, reliably produced cumulative semantic interference effects (Simulation 5), even without competitive lexical selection (Simulation 6). Thus, competitive selection was not needed for cumulative semantic interference when weight decreases occurred during learning. In fact, Simulation 6 demonstrated that, although the selection mechanism employed in the first five simulations was functionally non-competitive, the model still exhibited cumulative semantic interference.

Competitive lexical selection did, however, allow the model to generate cumulative semantic interference via weight strengthening alone (Simulation 6), as in Howard *et al.* (2006), by effectively converting an occlusion process into an inhibitory one (*cf.* Mensink & Raaijmakers, 1989). Thus cumulative semantic interference can derive from either *competitive selection*, where strong competitors interfere with target retrieval, or *competitive learning*, where strengthening a target involves weakening competitors. Therefore, we can derive a more general principle that lexical competition affects lexical selection, without constraining the point at which this competition comes into play. And error-based learning provides sufficient competition to explain cumulative semantic interference.

Our conclusion that competitive selection is not required to explain cumulative semantic interference has ramifications for the current debate about the necessity of such a selection process. It is fair to say that a majority of production researchers hold that lexical selection is competitive; that is, activated competitors retard target retrieval (*e.g.* Levelt *et al.*, 1999). Competitive selection was thought to have been demonstrated in the picture-word interference paradigm, in which a seen or heard distractor item slows the naming of a picture presented at about the same time (Schriefers *et al.*, 1990). Mahon *et al.* (2007), however, presented findings suggesting that the influence of external distractors on naming response times occurs post-lexically, and so the relevance of semantic interference from this paradigm for competitive lexical selection can be questioned (see, *e.g.* Abdel Rahman & Melinger, 2009; Janssen, Schirm, Mahon, & Caramazza, 2008; Mahon & Caramazza, 2009, for recent discussion). Thus, cumulative semantic interference, in which interference is generated from previous naming trials rather than external stimuli, might be seen as better evidence for

competitive lexical selection (Dell, Oppenheim, & Kittredge, 2007; Howard *et al.*, 2006; Navarrete, *et al.*, 2008; Schnur *et al.*, 2006). Here is where our modeling exercise matters. Although the model implicates a role for competition in explaining semantic interference, it does not require a competitive selection process to simulate the data. On the one hand, the model reinforces the conclusion of Howard *et al.*, (2006), that some kind of competition is required to explain cumulative semantic interference. On the other hand, the model offers the novel hypothesis that this competition arises through learning, rather than through a lexical selection mechanism that is slowed by any activated competitor. If this hypothesis is true (as well as concerns about the relevance of the picture-word paradigm), then we are back at square one on the question of competitive lexical selection.

### Incremental learning as an account of cumulative semantic interference

The model uses the delta rule, an error-based learning algorithm, to explain cumulative semantic interference. The important aspect of this algorithm is that it creates both strengthening of the connections to the target, and the weakening of the connections to competitors. Connection strengthening is clearly required for repetition priming, and connection weakening appears to explain at least a component of semantic interference and the strength of the perseveratory lag effect. Moreover, connection weakening is necessary for semantic interference if it is assumed that lexical selection is not competitive. Thus, the delta rule motivates an account of the data that combines the two principal hypothesized explanations for semantic interference and retrieval induced forgetting in general—the occlusion and the inhibition hypotheses. The fact that the weakening and strengthening is directly proportional to error, is not, as far as we can tell, directly relevant to the model's account of the data. But this aspect of the delta rule is motivated by research on Pavlovian conditioning (*e.g.* Rescorla & Wagner, 1972), episodic memory encoding (*e.g.* McClelland, McNaughton, & O'Reilly, 1995), frequency sensitivity in priming (*e.g.* Chang *et al.*, 2006), greater repetition priming in more error-prone conditions (*e.g.* Anderson, 2008), and division of labor effects in multi-component computational systems (*e.g.* Harm & Seidenberg, 2004).

### Extending and testing the model

Though not formally simulated, the model is also compatible with several other empirically-established effects of cumulative semantic interference on lexical retrieval times:

1. **Interference is robust to timing manipulations** (*e.g.* intervening non-verbal fillers: Damian & Als, 2005, Experiment 1; RSI manipulations: Hsiao *et al.*, 2009; Schnur *et al.*, 2006, Experiments 1 & 2; *cf.* simultaneous presentation: Belke *et al.*, 2005, Experiment 2). In our model, cumulative semantic interference derives from incremental learning and therefore follows the same time course as the learning process itself. That is, it persists without regard to time.
2. **Interference is robust to filler material in other paradigms** (*e.g.* naming pictures from other categories in the blocked-cyclic paradigm: Damian & Als, 2005, Experiments 2–4). In line with Simulation 1, this robustness comes from the fact that only relevant experience leads to relevant learning, and hence priming or interference. As long as filler material is sufficiently orthogonal to the critical items, it should never affect the build-up or resolution of semantic interference.
3. **Interference effects are graded as a function of semantic similarity** (Vigliocco *et al.*, 2002). In other words, more similar sets of items produce more interference than less similar sets. Such graded effects naturally emerge from the use of distributed semantic representations, where similarity is a function of feature overlap instead of discrete category membership.

4. **Interference is task-dependent** (*e.g.* Damian *et al.*, 2001). It requires mapping from shared semantic representations to separate lexical representations. Therefore, tasks that engage either type of representation, but involve no such mapping, should not elicit interference. For instance, non-verbally categorizing pictures according to visual or semantic features (*e.g.* Damian *et al.*, 2001) should not involve the semantic-to-lexical mapping, and should therefore not show our cumulative semantic interference. Similarly, orthographically cued word naming (*e.g.* Damian *et al.*, 2001, Experiment 2; Kroll & Stewart, 1994, Experiment 1) should only elicit semantic interference to the extent that utterance planning requires semantic access.
5. **Interference is cue-independent.** For example, Wheeldon and Monsell, (1994) used naming-to-definition to prime picture-naming, suggesting that the prime affected mappings from amodal semantic representations to lexical items. Our model is consistent with this cue-independence because semantic interference effects are carried in the semantic-to-lexical connections. Any process that uses these connections should therefore show cumulative semantic interference, regardless of the instigating stimulus.

These five response time effects derive from the fact that the model attributes semantic interference to incremental learning during lexical access, as opposed to some kind of time-dependent facilitatory or inhibitory priming. Indeed, these findings could also be readily explained by the model of Howard *et al.* (2006), which was the first implemented account of semantic interference in production based on persistent changes. We consider our model to be a descendant of that model. However, our model differs from its ancestor in three respects. First, it links up with speech-error based models of production and hence accounts for errors, including perseverations and omissions. This allows the model to simulate aphasic error data, and to offer a mechanism for how the brain chooses among active lexical candidates. Second, it ascribes an important role to connection weakening in explaining semantic interference effects on RT's and errors, and specifically the perseveration lag function. Connection weakening is a natural consequence of the delta rule, an algorithm for incremental learning. And third, the model represents the possibility that the competition needed to explain interference may not occur during lexical selection; it can arise from learning. As a result, non-competitive lexical selection becomes a viable account for data in this paradigm, and for production in general.

#### Additional predictions of our model

Our model offers additional predictions, mostly stemming from the idea that cumulative semantic interference reflects incremental learning. First, because the model's learning is error-based and not dependent on actual selection, cumulative semantic interference should accrue even when a name is not correctly retrieved, as in the case of errors and omissions. Although this property has not been tested for cumulative semantic interference, it appears to hold true for retrieval-induced forgetting. Using a part-set cueing paradigm, Storm, Bjork, Bjork and Nestojko (2006) demonstrated that providing misleading practice cues (*i.e.* they either elicited novel associations or offered no valid response) created retrieval-induced forgetting equivalent to that from valid cues. Therefore we expect, and our model predicts, that semantic interference should even accrue from naming trials that elicit omission errors.

Also, because cumulative semantic interference is attributed to persistent changes in connection weights, the effects should not spontaneously dissipate. Several picture-naming studies (*e.g.* Howard *et al.*, 2006; Nickels, Howard, Dodd, & Coltheart, 2008) have suggested that cumulative semantic interference persists at relatively long item-lags (minutes). But longer periods of persistence have not been examined in the picture-naming

paradigm. One episodic memory study (Anderson & Spellman, 1995) concluded that retrieval-induced forgetting persisted over lags of twenty minutes and another (Storm *et al.*, 2006) even reported that retrieval-induced forgetting effects may remain detectable after a 1-week lag (but see MacLeod & Macrae, 2001, and Postman *et al.*, 1968, for some evidence to the contrary). So, with the caveat that relevant experience is easier to define in a model than in the real world, we should find that cumulative semantic interference dissipates largely as a function of relevant experience, rather than as a function of time.

Finally, since our model derives from domain-general principles, it should formally extend to non-linguistic processes. Any system that involves shared activation, activation-dependent selection, and competitive learning should exhibit similar effects, and may be examined through similar experimental paradigms. Throughout this paper, we have referred to retrieval-induced forgetting (RIF), a well-known episodic memory effect, where the process of retrieving one association leads to impaired recall of competing associations. Several prominent accounts of RIF (*e.g.* Anderson, 2003; Norman *et al.*, 2007) suggest that it arises from the process of new information overwriting the old. In other words, they ascribe the memory effect to the process of learning. Such explanations have also surfaced for effects in visual object recognition (*e.g.* Marsolek, Schnyer, Deason, Ritchey, & Verfaellie, 2006). Thus, we can identify cumulative semantic interference with a more general theme in the way the mind operates, independent of the particular types of representations in use.

It may also be possible to extend the model's links with brain mechanisms and regions. The successful resolution of semantic interference may depend on processes subserved by the LIFG and/or left temporal lobe (LT) and thus may be impaired by damage to these regions. Evidence for this localization comes from both neuroimaging studies of healthy subjects (Maess, Friederici, Damian, Meyer, & Levelt, 2002; Moss *et al.*, 2005, Schnur *et al.*, 2009; see also Hocking, McMahon, & de Zubicaray, 2008) and lesion-mapping studies in patients (Schnur *et al.*, 2005, 2009). The involvement of frontal processes is further supported by single-case and group studies that associate exaggerated blocking effects with nonfluency and other symptoms of anterior aphasia (Biegler, Crowther & Martin, 2008; McCarthy & Kartsounis, 2000; Schnur *et al.*, 2006; Wilshire & McCarthy, 2002). Consistent with the LT localization, Simulation 4 showed that adding noise to lexical activations succeeded in reproducing the aphasic blocking pattern. In a follow-up test (not reported in this article), we simulated LIFG damage by instead adding noise to the booster mechanism and found that this gave the same result as the simulated lexical damage. While these simulations of brain damage are encouraging, we note that they fail to capture some subtle but potentially important distinctions. For one thing, it seems that while lesions to either the LIFG or LT exaggerate the blocking effect, only LIFG lesions produce the pattern of increased errors across cycles (Schnur *et al.*, 2005; 2006); in our simulations, noisy lexical activations and a noisy booster both had this effect. Also, whereas our simulations of aphasic damage focused on errors (in keeping with the data from Schnur *et al.*, 2006), recent evidence suggests that some frontal aphasics manifest exaggerated blocking interference in naming latencies rather than errors (Biegler *et al.*, 2008). To resolve these issues, it may be useful to measure the time course of LIFG and LT activation with imaging methods with good temporal resolution (*e.g.* the EROS optical imaging technique, Tse *et al.*, 2007). If interference in, say, the blocking paradigm is present in lexical-semantic areas (LT), but resolved in the LIFG, then manipulations of degree of interference should manifest first in the former region, and then the latter.

## Conclusion

The model instantiates a dynamic view of lexical knowledge. Shared semantic representations put competing words in a dynamic equilibrium where no semantic feature connects too strongly to any one word. Each act of lexical retrieval produces persistent, competitive, learning that perturbs this balance. It facilitates repeating the same word and impairs access to competing words. But retrieving a competitor shifts the balance back again. So not only are we capable of learning new words every day, but we are constantly adjusting which words, of the ones we know, are more or less available for use in speaking.

We believe that our model offers a parsimonious account of the lexical production data dealing with persistent effects, and it has implications for other issues in production, learning, and memory. Similarity-based interference that results from learning is not unique to lexical retrieval (*e.g.* Anderson, 2003; Marsolek *et al.*, 2006). But, when words are retrieved in a semantically-manipulated context, the resulting impairment is what we know as cumulative semantic interference.

## Acknowledgments

This research was supported by National Institutes of Health Grants DC000191, HD44458, and MH1819990. We thank Aaron Benjamin, Kara Federmeier, Matt Goldrick, John Hummel, Brad Mahon, Randi Martin, Sharon Thompson-Schill, Tatiana Schnur, and the Bock and Dell joint lab meeting for comments on this work and other contributions.

## References

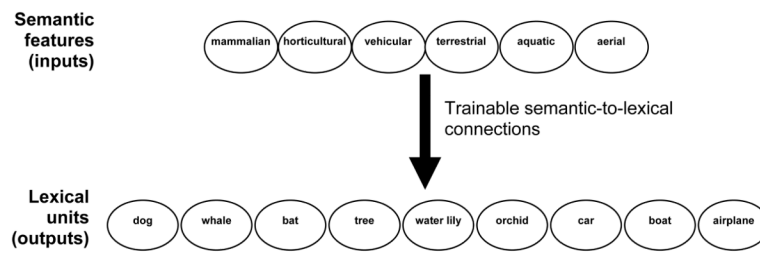
- Abdel Rahman R, Melinger A. When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2007;33(3):604–614.
- Anderson JD. Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of Fluency Disorders* 2008;33(2):135–155. [PubMed: 18617053]
- Anderson MC, Bjork RA, Bjork EL. Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1994;20(5):1063–1087.
- Anderson, MC.; Neely, JH. Interference and inhibition in memory retrieval. In: Bjork, EL.; Bjork, RA., editors. *Memory*. San Diego, CA: Academic Press; 1996. p. 237-313.
- Anderson MC, Spellman BA. On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review* 1995;102(1):68–100. [PubMed: 7878163]
- Anderson MC. Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language* 2003;49(4):415–445.
- Belke E. Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review* 2008;15(2):357–363. [PubMed: 18488652]
- Belke E, Meyer AS, Damian MF. Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology* 2005;58(4):667–692.
- Biegler KA, Crowther JE, Martin RC. Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology* 2008;25:493–527. [PubMed: 19086200]
- Blaxton TA, Neely JH. Inhibition from semantically related primes: Evidence of a category-specific inhibition. *Memory & Cognition* 1983;11(5):500–510.
- Brown AS. Priming effects in semantic memory retrieval processes. *Journal of Experimental Psychology: Human Learning and Memory* 1979;5(2):65–77.
- Bock K, Griffin ZM. The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General* 2000;129(2):177–192. [PubMed: 10868333]

- Brown AS. Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory* 1981;7(3):204–215.
- Castro, YG.; Strijkers, K.; Costa, A.; Alario, X. When cat competes with dog but not with perro: Evidence from the semantic competitor paradigm. Poster presented at the Fifth International Workshop on Language Production; Annapolis, MD. 2008 Jul.
- Chang F, Dell GS, Bock K. Becoming syntactic. *Psychological Review* 2006;113(2):234–272. [PubMed: 16637761]
- Cohen L, Dehaene S. Competition between past and present: Assessment and interpretation of verbal perseverations. *Brain* 1998;121:1641–1659. [PubMed: 9762954]
- Collins AM, Loftus EF. A spreading activation theory of semantic memory. *Psychological Review* 1975;82:407–428.
- Collins AM, Quillian ER. Retrieval time from semantic memory. *Journal of verbal learning and verbal memory* 1969;8:240–247.
- Crowther, JE.; Martin, RC.; Biegler, KA. The semantic blocking effect in naming: A computational model of inhibition failure. Poster presented at the 46th Annual Meeting of the Academy of Aphasia; Turku, Finland. 2008 Oct.
- Damian MF, Als LC. Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2005;31(6):1372–1384.
- Damian MF, Vigliocco G, Levelt WJ. Effects of semantic context in the naming of pictures and words. *Cognition* 2001;81(3):B77–86. [PubMed: 11483172]
- Dell GS. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 1986;93(3):283–321. [PubMed: 3749399]
- Dell GS, Burger LK, Svec WR. Language production and serial order: A functional analysis and a model. *Psychological Review* 1997;104(1):123–147. [PubMed: 9009882]
- Dell GS, Oppenheim GM, Kittredge AK. Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes* 2008;23:583–608. [PubMed: 20622975]
- Forde EME, Humphreys GW. A semantic locus for refractory behaviour: Implications for access-storage distinctions and the nature of semantic memory. *Cognitive Neuropsychology* 1997;14(3):367–402.
- Goldinger SD. Echoes of echoes? an episodic theory of lexical access. *Psychological Review* 1998;105(2):251–279. [PubMed: 9577239]
- Gordon JK, Dell GS. Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science* 2003;27:1–40.
- Gupta P, Cohen NJ. Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review* 2002;109(2):401–448. [PubMed: 11990324]
- Harm MW, Seidenberg MS. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review* 2004;111(3):662–720. [PubMed: 15250780]
- Hocking J, McMahon KL, de Zubicaray GI. Semantic context and visual feature effects in object naming: An fMRI study using arterial spin labeling. *Journal of Cognitive Neuroscience*. 2008
- Houghton, G. The problem of serial order: A neural network model of sequence learning and recall. In: Dale, R.; Mellish, C.; Zock, M., editors. *Current research in natural language generation*. London: Academic Press; 1990. p. 287-319.
- Howard D, Nickels L, Coltheart M, Cole-Virtue J. Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition* 2006;100(3):464–482. [PubMed: 16413014]
- Hsiao EY, Schwartz MF, Schnur TT, Dell GS. Temporal characteristics of semantic perseverations induced by blocked-cyclic picture naming. *Brain and Language* 2009;108:133–144. [PubMed: 19138794]
- Janssen N, Schirm W, Mahon BZ, Caramazza A. Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2008;34(1):249–256.

- Kan IP, Thompson-Schill SL. Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective & Behavioral Neuroscience* 2004;4(1):43–57.
- Kraljic T, Samuel AG. Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology* 2005;51(2):141–178. [PubMed: 16095588]
- Kroll JF, Stewart E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 1994;33:149–174.
- Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 1999;22(1):1–38. [PubMed: 11301520]
- MacLeod, CM.; Dodd, MD.; Sheard, ED.; Wilson, DE.; Bibi, U. In opposition to inhibition. In: Ross, BH., editor. *The psychology of learning and motivation*. San Diego, CA: Academic Press; 2003. p. 163-214.
- MacLeod MD, Macrae CN. Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science : A Journal of the American Psychological Society/APS* 2001;12(2):148–152. [PubMed: 11340924]
- Maess B, Friederici AD, Damian M, Meyer AS, Levelt WJM. Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience* 2002;14(3):455–462. [PubMed: 11970804]
- Mahon BZ, Caramazza A. Why does lexical selection have to be so hard? comment on Abdel Rahman and Melinger's swinging lexical network proposal. *Language and Cognitive Processes* 2009;24(5):735.
- Mahon BZ, Costa A, Peterson R, Vargas KA, Caramazza A. 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* 2007;33(3):503–535.
- Marsolek CJ, Schnyer DM, Deason RG, Ritchey M, Verfaellie M. Visual antipriming: Evidence for ongoing adjustments of superimposed visual object representations. *Cognitive, Affective & Behavioral Neuroscience* 2006;6(3):163–174.
- McCarthy RA, Kartsounis LD. Wobbly words: refractory anomia with preserved semantics. *Neurocase* 2000;6:487–97.
- McClelland JL, McNaughton BL, O'Reilly RC. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 1995;102(3):419–457. [PubMed: 7624455]
- McClelland JL, Rogers TT. The parallel distributed processing approach to semantic cognition. *Nature Reviews: Neuroscience* 2003;4(4):310–322.
- McGeoch JA. Forgetting and the law of disuse. *Psychological Review* 1932;39(4):352–370.
- Melton AW, Irwin JM. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *The American Journal of Psychology* 1940;100(3–4):610–641. [PubMed: 3322059]
- Mensink G, Raaijmakers JG. A model for interference and forgetting. *Psychological Review* 1988;95(4):434–455.
- Mitchell DB, Brown AS. Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1988;14(2):213–222.
- Moss HE, Abdallah S, Fletcher P, Bright P, Pilgrim L, Acres K, et al. Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. *Cerebral Cortex* 2005;15:1723–1735. [PubMed: 15728742]
- Navarrete, E.; Mahon, BZ.; Caramazza, A. The cumulative semantic cost in picture naming (*Il Costo Cumulativo Intracategoriale in compiti di denominazione di figure*). Paper presented at the XIV Congresso Nazionale Associazione Italiana di Psicologia; Padova, Italia. 2008 Sep.
- Nickels, L.; Howard, D.; Dodd, H.; Coltheart, M. Further investigations into cumulative semantic inhibition in picture naming: Effects of long lag & repeated members of a category. Poster presented at the Fifth International Workshop on Language Production; Annapolis, MD. 2008 Jul.

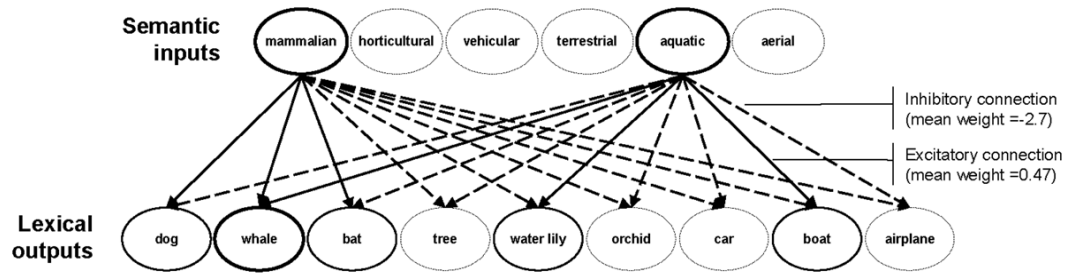


- Norman KA, Newman EL, Detre G. A neural network model of retrieval-induced forgetting. *Psychological Review* 2007;114(4):887–953. [PubMed: 17907868]
- Oppenheim GM, Dell GS, Schwartz MF. Cumulative semantic interference as learning. *Brain and Language* 2007;103:175–176.
- Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 1996;103(1):56–115. [PubMed: 8650300]
- Postman L, Stark K, Fraser J. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior* 1968;7:672–694.
- Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. *Psychological Review* 2000;107(3):460–499. [PubMed: 10941277]
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts; 1972. p. 64-99.
- Roediger HL, Neely JH, Blaxton TA. Inhibition from related primes in semantic memory retrieval: A reappraisal of brown's (1979) paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1983;9(3):478–485.
- Roelofs A. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 1992;42(1–3):107–142. [PubMed: 1582154]
- Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, et al. Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review* 2004;111(1):205–235. [PubMed: 14756594]
- Rumelhart, DE.; McClelland, JL. the PDP research group. *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. I. Cambridge, MA: MIT Press; 1986.
- Schnur TT, Lee E, Coslett HB, Schwartz MF, Thompson-Schill SL. When lexical selection gets tough, the LIFG gets going: A lesion analysis study of interference during word production. *Brain and Language* 2005;95(1):12 – 13.
- Schnur TT, Schwartz MF, Brecher A, Hodgson C. Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language* 2006;54(2):199–227.
- Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences of the United States of America*. (in press).
- Schriefers H, Meyer AS, Levelt WJM. Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language* 1990;29:86–102.
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America* 1997;94(26):14792–14797. [PubMed: 9405692]
- Tse CY, Lee CL, Sullivan J, Garnsey SM, Dell GS, Fabiani M, et al. Imaging cortical dynamics of language processing with the event-related optical signal. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104(43):17157–17162. [PubMed: 17942677]
- Vigliocco G, Vinson DP, Damian MF, Levelt WJM. Semantic distance effects on object and action naming. *Cognition* 2002;85(3):B61–9. [PubMed: 12169413]
- Wheeldon LR, Monsell S. Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language* 1994;33:332–356.
- Widrow, B.; Hoff, ME. IRE WESCON convention record. New York: IRE; 1960. Adaptive switching circuits; p. 96-104.(Reprinted in Anderson, J. A., & Rosenfeld, E. (Eds.) (1988). *Neurocomputing: Foundations of research* (pp. 123–134). Cambridge, MA: MIT Press.)
- Wilshire CE, McCarthy RA. Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology* 2002;19:165–186.



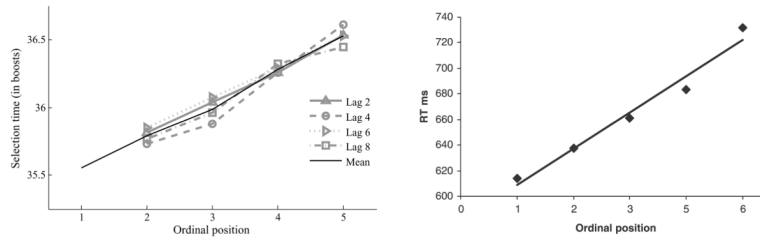
**Figure 1.**

Words are activated by distributed semantic representations, or features. The strength of a connection from a semantic feature to a word is established by an initial period of error-based learning, and this learning continues throughout the simulations.



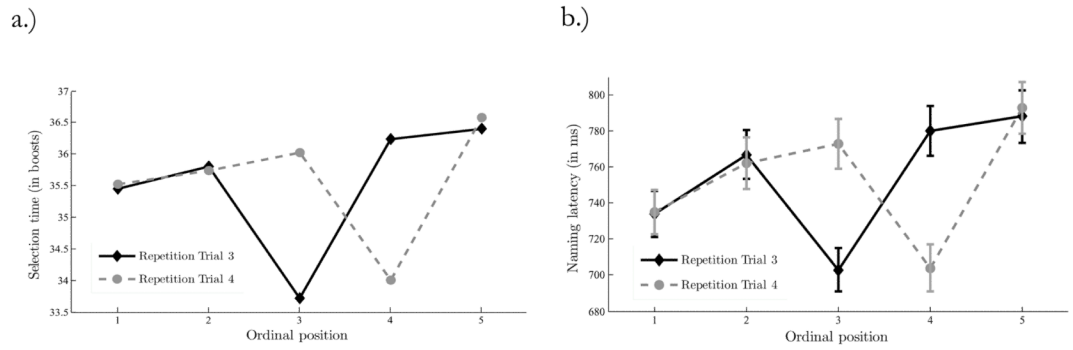
**Figure 2.**

A scaled-down illustration of a trained network's vocabulary. Here, the feature MAMMALIAN excites three words (DOG, WHALE, and BAT) and inhibits the rest. The feature AQUATIC works the same way. When these features are activated together, they activate all of the words that are either mammalian or aquatic. But they activate WHALE most strongly because it is both mammalian *and* aquatic.

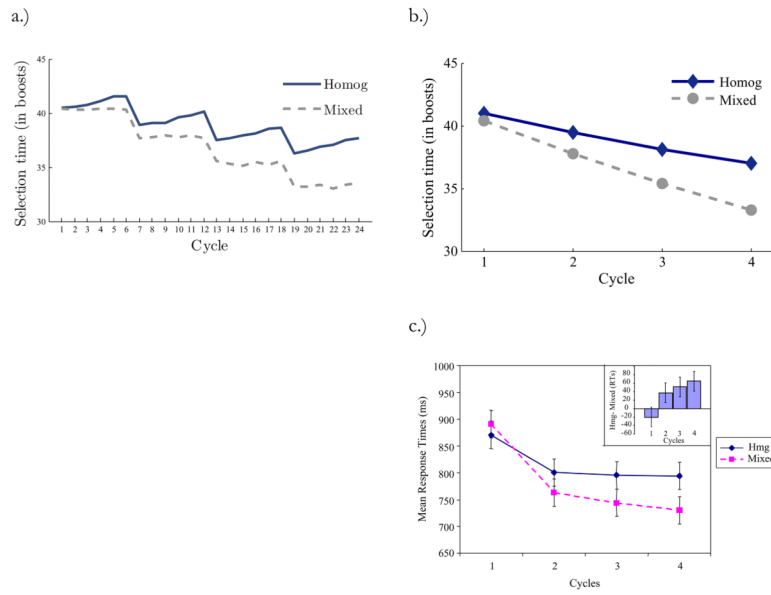


**Figure 3.**

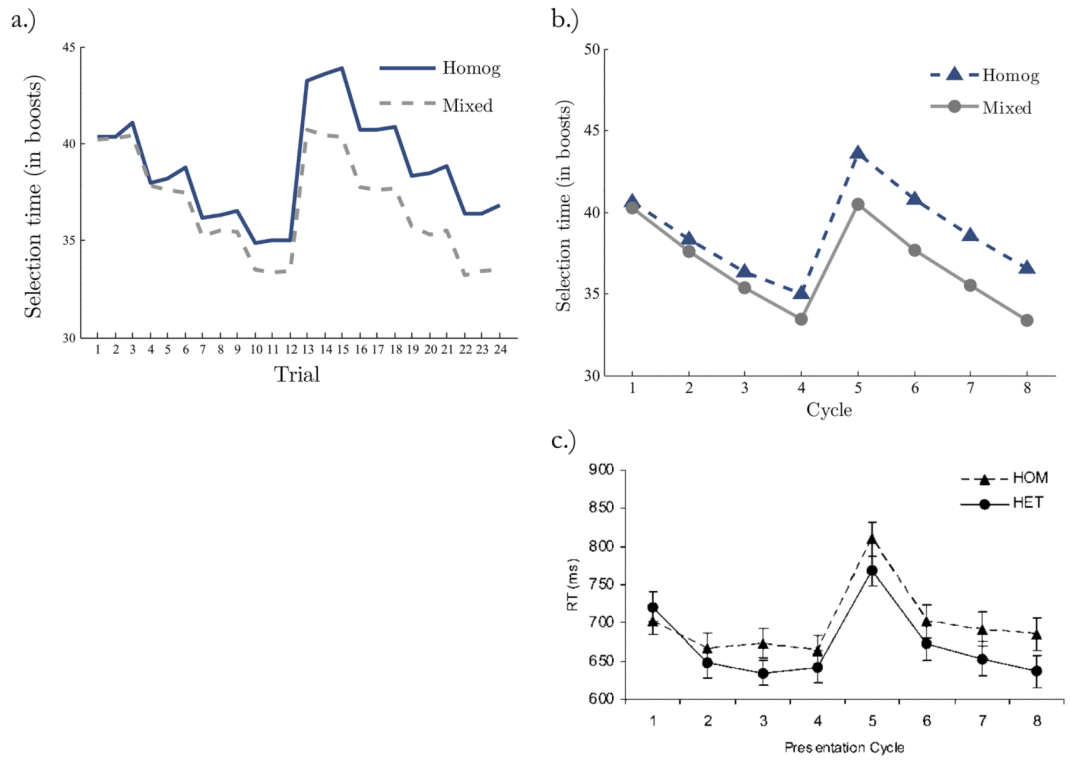
In Simulation 1a, the continuous paradigm, each critical item takes a bit longer to select than the previous one, regardless of the number of intervening unrelated items. a.) Model-predicted mean correct selection times in each of the four lag conditions show identical increases with ordinal position. We present only an overall mean selection time for ordinal position 1 because the lag for the first item is obviously impossible to calculate. b.) Human subjects' mean naming latencies show the same increase (reprinted from Howard *et al.*, 2006).



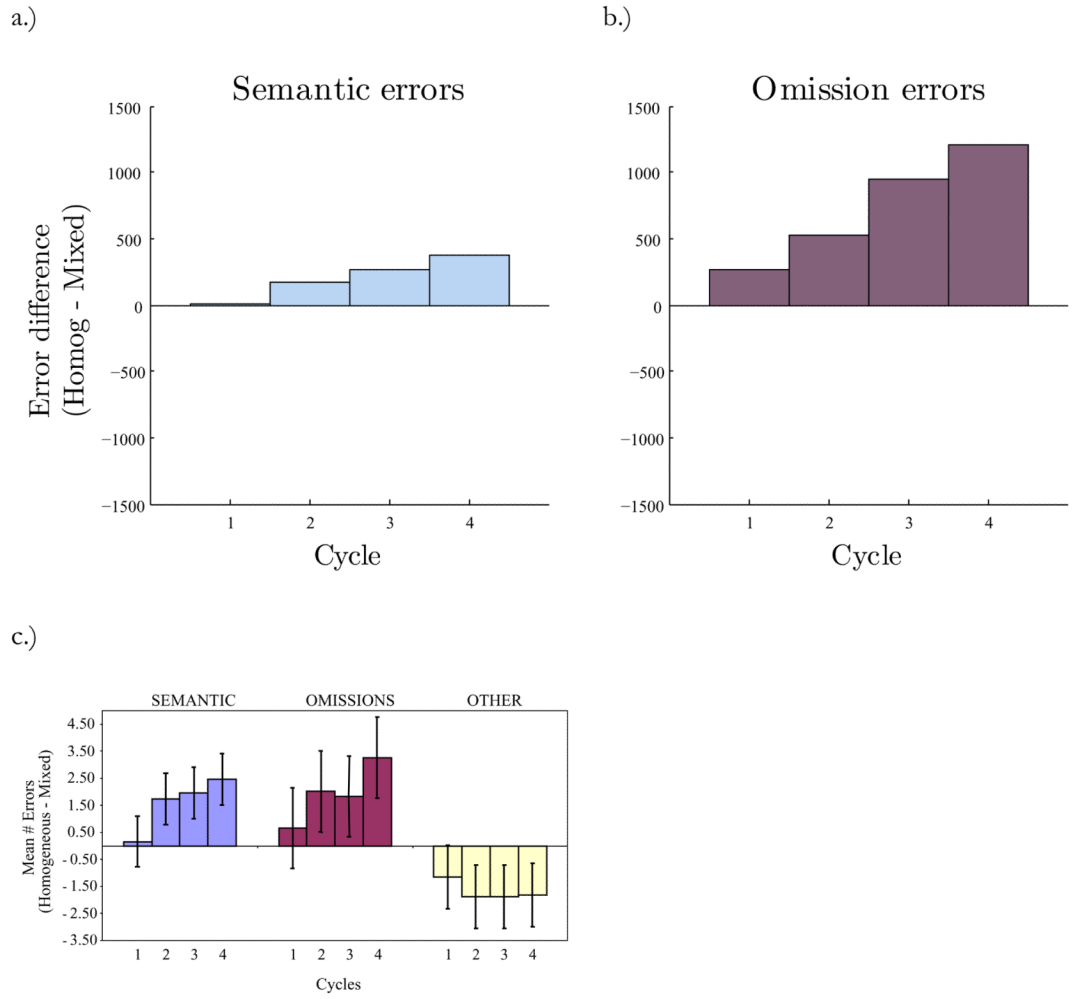
**Figure 4.** Repeating an item creates cumulative semantic interference similar to that from accessing a novel item. a.) Model-predicted mean selection times from Simulation 1b. b.) Empirical results from Navarrete *et al.* (2008).



**Figure 5.** Simulated lexical selection times from Simulation 2 mirror human subjects' naming latency patterns for the blocked-cyclic paradigm. a.) Mean simulated lexical selection times across four cycles through sets of six words, for 10,000 networks. b.) Predicted per-cycle means, derived from (a), for comparison with the human data presented in (c). c.) Subjects' per-cycle mean naming latencies across four cycles, from Schnur *et al.* (2006).

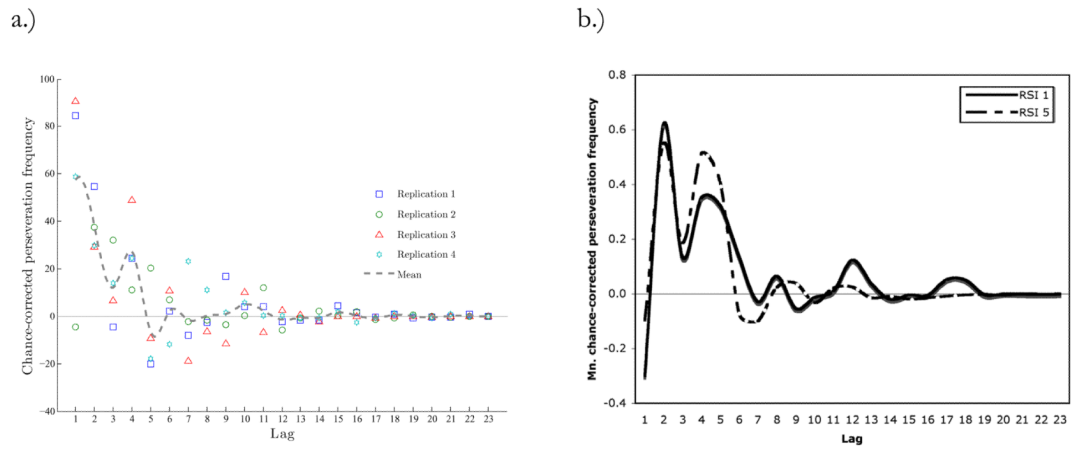


**Figure 6.** Results from Simulation 3. a.) Lexical selection times for items in homogeneous and mixed sets presented in three-trial cycles. The second subset replaced the first after the twelfth trial. b.) Per-cycle mean selection latencies, derived from (a), for comparison with (c). c.) Human subjects' per-cycle mean naming latencies, from Belke *et al.* (2005).

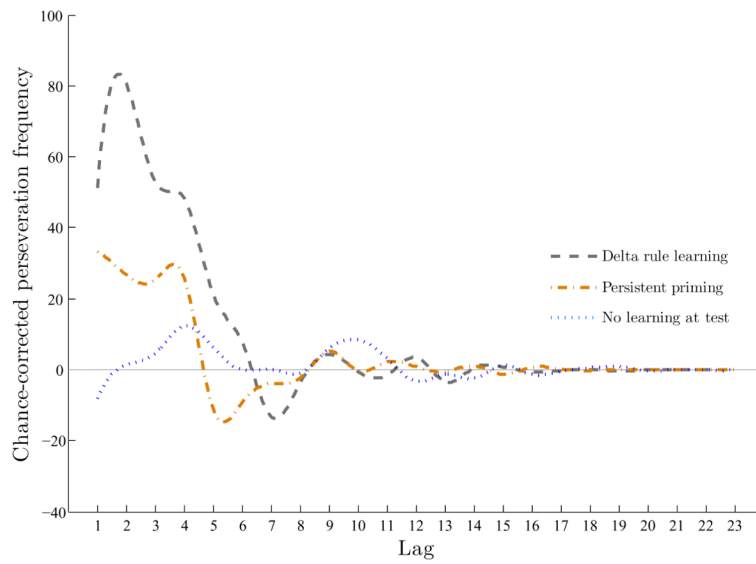


**Figure 7.** Predicted semantic error and omission patterns from Simulation 4, compared to patient error data for the same task. a.) Simulated semantic errors become increasingly more frequent in the Homogeneous condition, compared to the Mixed condition. b.) The differences in the omission error rates between the Homogeneous and Mixed conditions also increase across cycles. c.) Patient error data from Schnur *et al.* (2006).

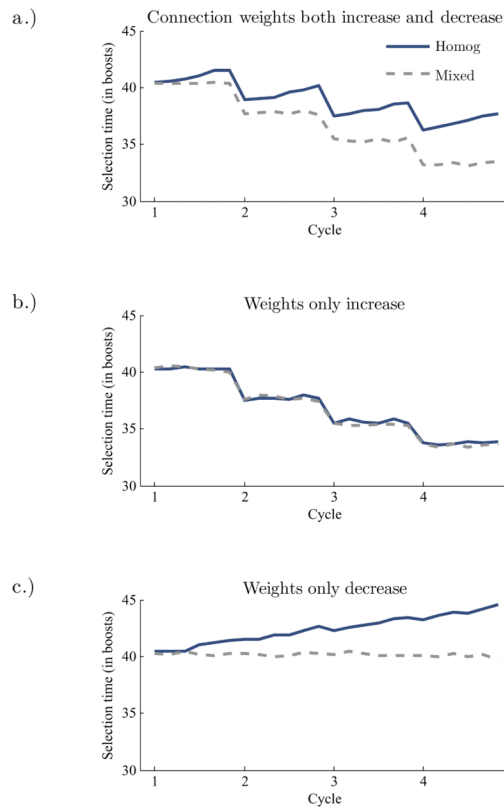




**Figure 8.** Chance-corrected perseveration frequencies in the homogeneous condition, plotted as a function of item-lag. a.) In Simulation 4, perseveration errors at short lags are more frequent than would be expected by chance. b.) Patients' perseverations, from Hsiao *et al.* (2009), are also more frequent at short lags.

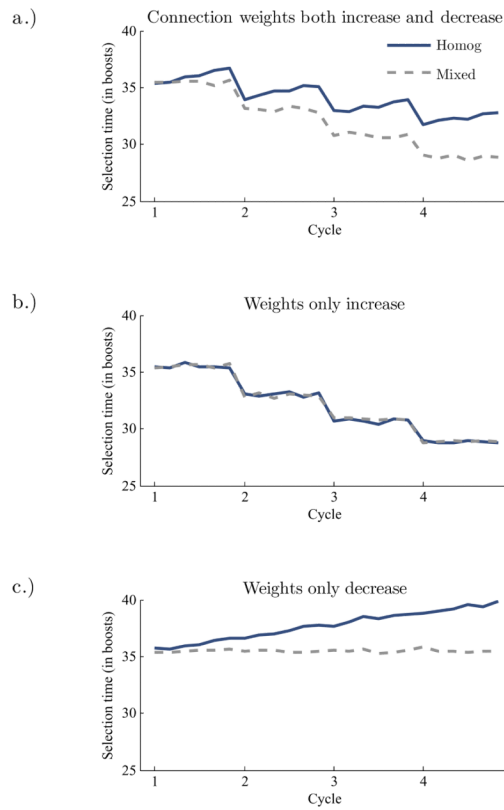


**Figure 9.** Chance-corrected perseveration frequencies as a function of item-lag, using three different learning rules during the testing phase: 1.) Error-based delta rule learning, 2.) Non-error based fixed-increment strengthening (priming), and 3.) No learning. As in Figure 8a, each spline represents a mean of four 10,000-block replications, though here we have omitted the individual data points in favor of visual clarity.

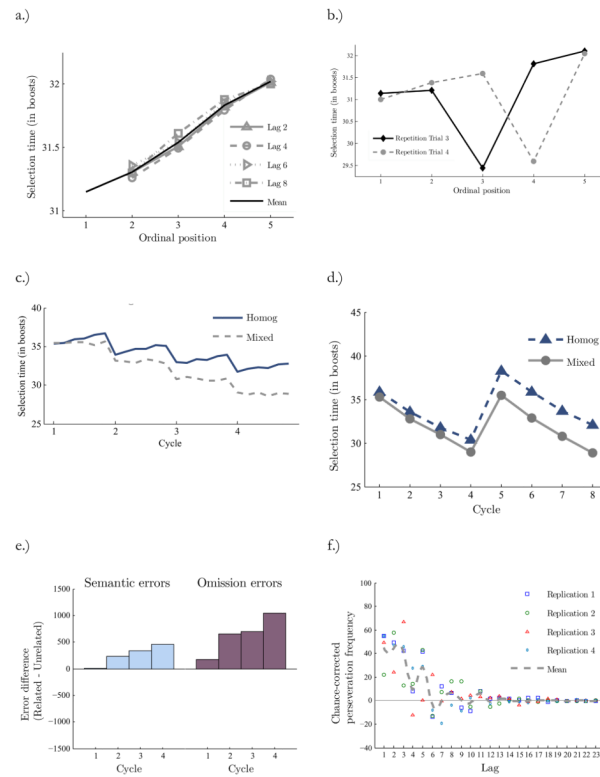


**Figure 10.**

Strengthening connections to target words produces a repetition priming effect, while weakening connections to competitors produces a semantic interference effect. a.) Applying both weight increases and decreases (Simulation 2) generates selection latencies resembling a saw-tooth pattern. b.) Increasing connection weights to target words (Simulation 5a) decreases selection times in each cycle, creating a step-like pattern. c.) Decreasing weights to competitors (Simulation 5b) continually increases selection times in the homogeneous condition.

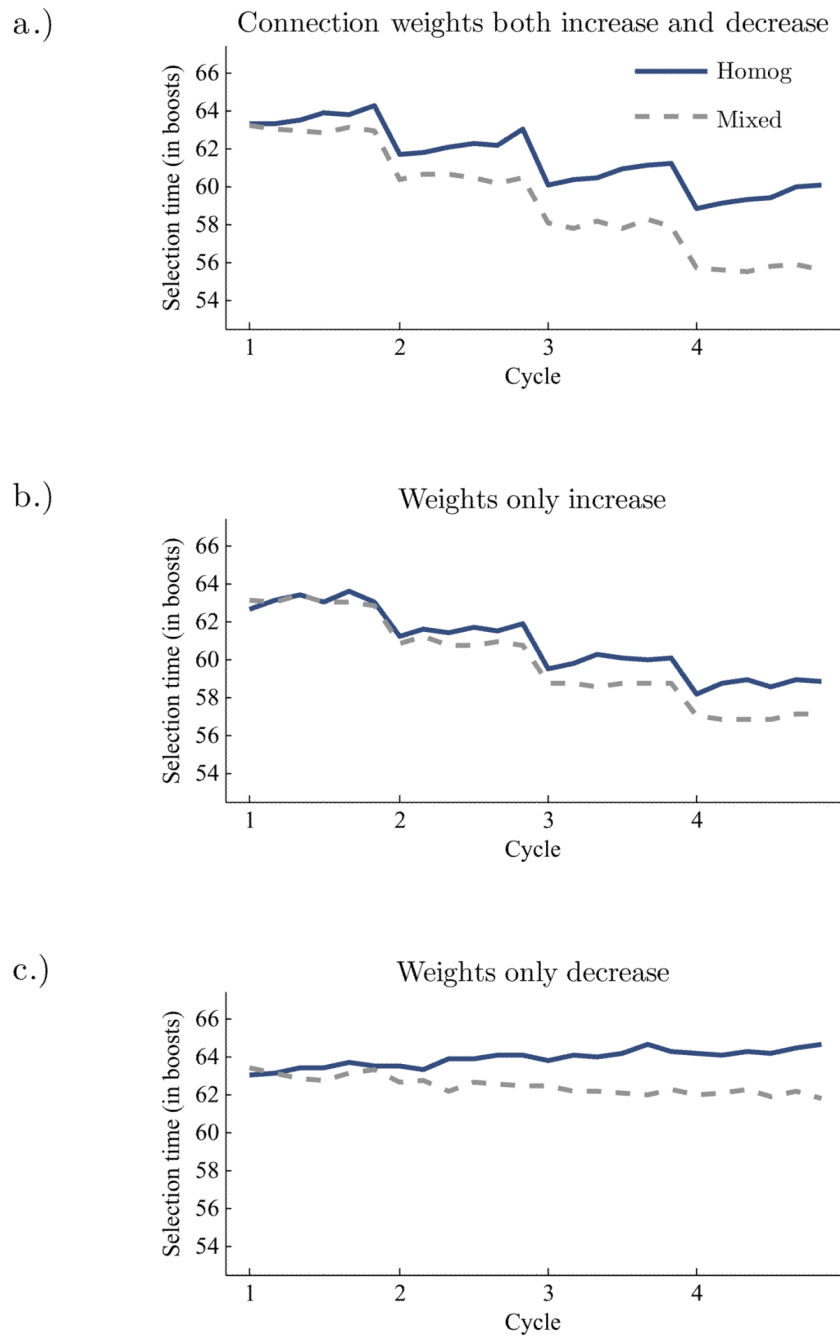


**Figure 11.** Predictions from the non-competitive selection model (Equation 9) mirror those from our competitive model in Simulation 5. a.) Implementing both weight increases and decreases produces the familiar saw-toothed pattern. b.) Weight increases alone produce the step function associated with repetition priming. c.) Weight decreases still lead to a cumulative semantic interference effect.



**Figure 12.**

Replicating Simulations 1–4 with a non-competitive selection mechanism gives results identical to those already presented. a.) Lag-insensitive increases in selection times as a function of ordinal position (Simulation 1a replication). b.) Repeated items also contribute to the interference effect (Simulation 1b replication). c.) Selection times increase within a cycle and decrease across cycles (Simulation 2 replication). d.) Accumulated interference generalizes to other same-category items (Simulation 3 replication). e.) Error effects increase across cycles (Simulation 4 replication). f.) More recently used words are more likely to be perseverated (Simulation 4 replication).



**Figure 13.**

A more competitive selection rule (Equation 10) compares each word's activation with just its strongest competitor. a.) Applying both weight increases and decreases generates selection latencies resembling the saw-tooth pattern from Simulation 2. b.) Increasing connection weights to target words produces a less dramatic version of this sawtooth. c.) Decreasing weights to competitors continually increases selection times in the homogeneous condition while decreasing times in the mixed condition.

**Table 1**

Model parameters for Simulations 1, 2, 3, 5, and 6.

<b>Parameter</b>	<b>Value</b>
<i>Learning rate (<math>\eta</math>)</i>	0.75
<i>Activation noise (<math>\theta</math>)</i>	0.5
<i>Boosting rate (<math>\beta</math>)</i>	1.01
<i>Threshold (<math>\tau</math>)</i>	1
<i>Deadline (<math>\Omega</math>)</i>	100

**Table 2**

Model parameters for Simulation 4. Except for Activation noise, these parameters are identical to those given in Table 1.

Parameter	Value
<i>Learning rate (<math>\eta</math>)</i>	0.75
<i>Activation noise (<math>\theta</math>)</i>	1
<i>Boosting rate (<math>\beta</math>)</i>	1.01
<i>Threshold (<math>\tau</math>)</i>	1
<i>Deadline (<math>\Omega</math>)</i>	100