Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production

Gary S. Dell, **Gary M. Oppenheim**, and **Audrey K. Kittredge**
University of Illinois, Urbana-Champaign

Abstract

Retrieving a word in a sentence requires speakers to overcome syntagmatic, as well as paradigmatic interference. When accessing *cat* in "The cat chased the string," not only are similar competitors such as *dog* and *cap* activated, but also other words in the planned sentence, such as *chase* and *string*. We hypothesize that both types of interference impact the same stage of lexical access, and review connectionist models of production that use an error-driven learning algorithm to overcome that interference. This learning algorithm creates a mechanism that limits syntagmatic interference, the syntactic "traffic cop," a configuration of excitatory and inhibitory connections from syntactic-sequential states to lexical units. We relate the models to word and sentence production data, from both normal and aphasic speakers.

The last 20 years has seen an explosion of research on language production. Seminal experiments, such as Schriefers, Meyer, and Levelt's (1990) application of the picture-word interference paradigm to the time course of lexical access, have inspired hundreds of studies in this area. At the same time, investigations of impaired production have increasingly adopted a psycholinguistic (rather than linguistic or clinical) perspective, with the result that theories are now constrained by neuropsychological data, as well as experimental findings from unimpaired speakers (see, e.g. Caramazza, 1997; Martin, 2003; Nickels, 2001). Along with this empirical work, computational models have been devised and honed to the point that they offer quantitative accounts of much of the data (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Levelt, Roelofs, & Meyer, 1999; Rapp & Goldrick, 2000; Roelofs, 1997).

The bulk of this progress has concerned the production of single words rather than sentences. Although there are noteworthy exceptions to this generalization—studies of structural priming (e.g. Ferreira, 2003), agreement (e.g. Alario & Caramazza, 2002; Eberhard, Cutting, & Bock, 2005), and the incremental nature of production (e.g. Griffin & Bock, 2000) come to mind—there is little doubt that our understanding of how we name a picture of, say, a cat, far outstrips that of how we describe a picture as "The cat chases the string." The goal of this paper is to narrow this gap.

Specifically, we discuss an issue that confronts theorists who wish to extend accounts of single-word retrieval to sentences. When accessing a particular word in a sentence, the production system has to deal with interference from other words in the sentence, as well as interference from words that are similar to the current target. In single-word access, the only challenge for the speaker is to activate the lexical representations of the target, for example, "cat," rather than those of "dog" and "cap." Semantic and phonological competitors to a target are *paradigmatic* sources of interference, and current models provide fair accounts of

Gary S. Dell, Beckman Institute, University of Illinois, 405 N. Mathews Ave. Urbana IL 61820 USA, gdell@uiuc.edu, 217-244-1294 fax: 217- 333-2922.

how performance reflects this kind of interference (e.g. Goldrick, 2006; Ruml, Caramazza, Capasso, & Miceli, 2005; Schwartz, Dell, Martin, Gahl, & Sobel, 2006). *Syntagmatic* interference arises when lexical access occurs in the context of other words that have been or are about to be produced.[1] So, in "The cat chases the string," "string" and "chase" are syntagmatic competitors when "cat" is accessed. We know much less about how these competitors are dealt with.

Our central claim is that the production system overcomes the syntagmatic and paradigmatic interference that it experiences when producing words in context through an adaptive implicit learning process that operates throughout the speaker's lifetime. We first explore the role of learning by focusing on the *syntactic category constraint* on speech errors, the tendency for erroneously substituted words to belong to the same category (e.g. noun) as the target word. We argue that the syntactic category constraint reflects a learned lexical selection process and we characterize this process using the sentence production model of Gordon and Dell (2003). This model uses a connectionist error-driven learning algorithm that is sensitive to paradigmatic sources of error (from semantic competitors) and syntagmatic sources of error (from other words in the sentence). The result of this sensitivity is a dynamic lexical-selection mechanism that we call the *syntactic traffic cop*. We then show that the traffic cop plays a role in explaining differences in patients' abilities to retrieve semantically light and heavy verbs. Finally, we consider how syntagmatic interference from previously retrieved words interacts with paradigmatic semantic interference when picture names are repeatedly retrieved from the same semantic category (e.g. Belke, Meyer, & Damian, 2005). Again, we show that error-driven learning is the key for explaining the data. In both our analysis of the syntactic traffic cop and our investigation of semantic interference from repeated naming, our proposed mechanisms will be visible in the excitatory and inhibitory connections that mediate lexical retrieval, connections whose weights are a product of learning.

The Syntactic Category Constraint

Nearly 30 years ago, Nooteboom (1969) reported that "a mistakenly selected word always or nearly always belongs to the same word class as the intended word," and hence, "the grammatical structure of the phrase under construction imposes imperative restrictions on the selection of words." Crucially, the syntactic category constraint impacts all word-error types, including word substitutions associated with paradigmatic interference such as semantic (1) and formal (2) substitutions, and word movement errors, which are caused by syntagmatic interference (3) (examples from Garrett, 1975; Fromkin, 1971).

1. **pot** → *cup* (nouns); **most**→ *least* (adverbs); **read**→*written* (verbs)

2. **pressure**→ *present* (nouns)

3. the **class** will be about discussing the **test** → will be about discussing the *class* (nouns)

The "imperative restrictions" on lexical selection revealed in the syntactic category constraint are key to linking word- and sentence-production theory. The constraint suggests that sentence production employs an abstract syntactic representation, for example, a structural frame with slots for open-class lexical items (e.g. Garrett, 1975). The frame slots are labeled by syntactic category such that they will only accept the insertion of a lexical item of the proper category. Notice that such a constraint on insertion reduces some, but not all, sources of syntagmatic interference when a sentence is produced. In "The cat chases the

---

[1]The terms *paradigmatic* and *syntagmatic* were first used in the context of activation-based theories of production by Lecours and L'hermitte (1969).

string," the access of *cat* is controlled by a frame slot labeled for nouns. This should greatly limit syntagmatic interference, effectively preventing insertion of the other words from the sentence into cat's slot, except for the only other noun, *string*.

What is the relation between the restrictions imposed by the syntactic category constraint on lexical access and other influences on access? We propose that syntagmatic and paradigmatic sources of competition operate at the same processing stage, a stage at which words are selected based jointly on their semantic and syntactic properties to fill a particular slot in the sentence (see Figure 1). This stage has been called, among other names, lemma selection (Levelt et al., 1999), L-level selection (Rapp & Goldrick, 2000), and word-level convergence of the sequential and semantic systems (Chang, Dell, & Bock, 2006). Crucially, it is a level at which words are represented holistically, not as a set of phonological units, and are activated incrementally, in sequence—if word *i* precedes word *j*, then word *i* is activated first (e.g. Dell, 1986;MacKay, 1982). We propose that, at this level, the target competes with semantically related words and with words that have previously been or are about to be spoken (perseveratory and anticipatory competitors, respectively). We also believe that phonologically-related words can compete with the target at this level (Schwartz et al., 2006), but any competition from words that are related solely through phonology appears to be weak (Rapp & Goldrick, 2000) and so we will not consider this factor here. Crucially, competitors of all types are much less able to interfere when they are not of the target syntactic category. To flesh out this proposal, we begin by reviewing the role of the syntactic category constraint in semantic errors in aphasia and in the picture-word interference task.

Semantic Word Substitutions

Semantic substitutions made by normal speakers nearly always obey the syntactic category constraint (Garrett, 1975). "Pot" might be replaced by "cup," but not by "pour". The obedience of semantic errors to syntactic constraints is thought to reflect an independent syntactic influence on lexical access, but this is not certain because semantic similarity tends to imply syntactic similarity (see Vigliocco, Vinson, Lewis, & Garrett, 2004). An important source of data on this question comes from examination of the syntactic category constraint in aphasic semantic substitutions. A study of these errors by Berndt et al. (1997) points to a damageable and possibly syntactic mechanism associated with the category constraint. Berndt et al. collected lexical retrieval errors from aphasic individuals who named pictures and videos of objects (noun targets) and actions (verb targets). We focus on four of her eight patients who had similar overall lexical retrieval performance (56%–75% correct), but whose semantic errors matched the syntactic category of the target to different extents. For example, semantic errors such as **road** (noun)→*driving* (verb) and **write** (verb)→ *letter* (noun) would be counted as non-matching. Some patients' errors consistently obeyed the syntactic category constraint (e.g. HF, 92% matching out of 56 errors; HY, 84% out of 66 errors), whereas others did so at rates close to chance adherence (e.g. FM, 61% matching out of 31 errors; LR, 35% matching out of 40 errors). Thus, the difference between the patients whose errors matched their targets most of the time, and those whose matching rates were close to chance, was quite dramatic. Although the ambiguity of error-category coding in English single-word retrieval (e.g. "jump" can be a noun) could have added considerable noise to these data, the differences among the patients are powerful enough to stand out against this noise. We suggest that at least some of the patients with many cross-category errors suffered from damage to a mechanism that enforces the syntactic-category constraint.

Three aspects of Berndt et al.'s data support this claim. First, the two patients cited above, whose errors were just about as likely to cross categories as to respect them, do not appear to have difficulties accessing meaning from spoken single words, suggesting that cross-category errors reflect syntactic problems, rather than central semantic confusion between

actions and objects. In a photo-word matching test with nouns and verbs, and in a video comprehension test with verbs, the patients with cross-category production errors had little difficulty (FM, 97% correct; LR, 92% correct), and certainly did not have more difficulty than the patients whose errors obeyed the constraint (HF, 94% correct; HY, 85% correct). On the assumption that central semantic difficulties would hurt word comprehension as well as production, these results argue against a semantic locus for the variability in cross-category errors.

The second relevant aspect of Berndt et al.'s data concerns the production of cross-category errors by these patients in grammatical *sentences*. The tendency to generate cross-category errors was largely eliminated in FM and LR when items were named as completions of a heard sentence such as "To put words on paper with pen is to …" (write), instead of being given in response to a picture, video, or heard definition. Berndt et al. attributed this change to the "greater syntactic constraints that the completion task places on the production of words from specific grammatical classes (p. 84)." It seems as if patients such as FM and LR have some residual ability to use grammatical constraints in lexical retrieval, and this ability is revealed when they hear an experimenter-produced grammatical context that provides category cues. For HF and HY, the constraint is apparent even in single-word production. Possibly, they can (internally or overtly) instantiate their own suitable "context," as it "It's a (noun)" or "She's (verb)ing" when retrieving single words.

The third piece of evidence also concerns sentence production. If FM and LR have difficulty using syntactic constraints in their lexical retrieval without external cues, in contrast to the other two patients, this may have consequences for their sentence production. Berndt et al. report that FM and LR, but not HF and HY, score in the agrammatic range with respect to their production of inflectional morphology in sentences (Saffran, Berndt, & Schwartz, 1989). Productive use of inflectional morphology depends heavily on syntactic structure, and production models often associate inflection with categorically labeled syntactic frames (e.g. Dell, 1986; Garrett, 1975). In summary, these findings are suggestive of a damageable cognitive mechanism that enforces the syntactic category constraint.

Semantic and Syntactic Distractors in the Picture-Word Interference Task

In the picture-word interference task, speakers are slower to name pictures when a semantic distractor word is heard or seen at about the same time that the picture is presented (e.g. Schriefers et al., 1990). This effect appears to reflect the same kind of paradigmatic interference that underlies semantic substitution errors. Thus, we can ask whether something analogous to the syntactic category constraint operates in this task. Pechmann and Zerbst (2002) demonstrated such an effect. Participants had to name pictured objects (e.g. an apple) using a definite determiner noun phrase. Under these conditions, noun distractors led to slower naming times than non-noun distractors. Using the same task, Vigliocco, Vinson, and Siri (2005) simultaneously manipulated syntactic and semantic sources of interference. Their participants had to name pictures of actions with an inflected form of a verb. Distractors could be either in the same or different syntactic category as the target (verbs or nouns, respectively) and were either semantically close or distant to the target. Both syntactic and semantic interference was found. Semantically close distractors increased naming times by 37 ms compared to distant distractors, and same-category (verb) distractors created 36 ms more interference than noun distractors. The data are generally supportive of our assumption that syntactic category is limiting the competition from wrong-category competitors.

An important aspect of both the Pechmann and Zerbst (2002) and the Vigliocco et al. (2005) studies is that the syntactic-category effect was only observed when the pictures were named in a phrasal context. In Pechmann and Zerbst's experiments, naming objects with bare nouns rather than determiner-noun phrases did not lead to interference from noun distractors. In

Vigliocco et al., naming actions with a verb in citation form, rather than with an inflected verb, eliminated the syntactic influences. This restriction of syntactic effects to phrasal contexts offers further evidence that syntactic category constraints on lexical access are a specific mechanism of *sentence* production. It does, however, raise a question about why the semantic errors of some of Berndt et al.'s aphasic patients consistently obeyed the category constraint, even though the task was single-word naming. As we suggested earlier, these patients may be (sometimes internally) generating phrasal or inflectional contexts. For example, Berndt et al.'s patients produced inflected verbs sometimes (e.g. "driving"), although there was no requirement to do so. An alternative explanation is that there is something about the picture-word interference task that suppresses the expression of the category constraint in single-word contexts. This task is not as well understood as one might hope (e.g. Mahon, Costa, Vargas, & Caramazza, 2007).

The Syntactic Traffic Cop in a Learning-based Model of Sentence Production

We have hypothesized a mechanism that helps to limit interference during lexical access in phrasal contexts to words from the same syntactic category as the target word. Errors seem to be largely restricted to substitutions in the same category, and distractor words that do not belong to the target category create less interference in the picture-word interference task. But we have not yet characterized this mechanism's operation or origin. To do so, we turn to a model of production by Gordon and Dell (2003), called the *division of labor* (DoL) model.

The DoL model is a connectionist network that learns to map from a static representation of a sentence's meaning and a changing syntactic-sequential state to a sequence of words (Figure 2). The network's input layer contains semantic and syntactic units, and the output layer has localist word units. Connections run from input to output, and their strength, including whether they are excitatory or inhibitory, is set by an error-driven learning algorithm, the delta rule (e.g. Widrow & Hoff, 1988).

The activation pattern across the semantic units represents the message to be conveyed. For example, for the sentence, "The bird flies," features corresponding to the BIRD concept (*winged* and *animal*) and the FLYING concept (*motion* and *through_air*) would be activated. The semantic activation pattern remains fixed throughout the production of the sentence, thus representing the model's assumption of a static message. The activation of the syntactic units, in contrast, changes throughout the sentence. Syntactic units represent syntactically defined sequential states such as "beginning of the sentence" or "head of the subject NP." In Gordon and Dell's bare-bones implementation, all of the sentences that the model experienced exhibited the structure ***The* NOUN(sing) VERB***s*, and so only three such states, labeled *Det*, *Noun*, and *Verb* were required to distinguish the three positions in the sentence. At the beginning of production, the *Det* state is activated and the other states are inactive. After the determiner has been encoded, *Det* is turned off, and *Noun* becomes active, and so on until the end of the sentence. The activation from the semantic and syntactic units passed through weighted connections to the output word layer, and after the model had experienced sufficient training, the activation would converge on a sequence of single-word units, for example, first *the*, then *bird*, and finally *fly* for the message given above. It should be noted that the changes from one syntactic-sequential state to another are just built in to the program running the model. Thus, the implementation does not learn the states or the mechanism for cycling through them. Later, we will suggest that these things are learnable, as demonstrated in a related production model, the Dual-path model of Chang et al. (2006).

The connection weights are set by the model attempting to produce sentences, one word at a time. To the extent that the model's output is incorrect (e.g. the correct pattern might be full activation of *bird*, and no activation for any other word at a particular point in the sentence), the connection weights are changed according to the delta rule. In essence, the rule causes weights from active semantic (e.g. *winged*, *animal*, *motion*, *through_air*) and syntactic (e.g. *Noun*) units to be increased to the target lexical unit (e.g. *bird*), and decreased to all other erroneously active units.

To activate the correct word at the correct time, the model has to learn to overcome both paradigmatic and syntagmatic interference. When *bird* is the target, words that share semantic features with *bird* have a tendency to be more active than those that do not share features. For example, in Gordon and Dell's implementation, the feature *winged* is as predictive of *plane* as it is of *bird*, and so *plane* functions as a semantic competitor during the access of *bird*. The model's learning algorithm limits this paradigmatic interference, though. To the extent that there is an erroneous tendency to activate *plane* when *bird* is intended, the network grows inhibitory connections from the features of *bird* not shared with *plane*, such as *animal*, to the lexical unit for *plane*. When the network has been thoroughly trained, semantically similar competitors present little problem, unless the network is damaged in some way, or the activation of the competitor is artificially enhanced as it is in the picture-word interference paradigm.

The model's approach to limiting syntagmatic interference is less straightforward, but more interesting. Recall that the model's semantic input is static and encompasses the entire clause to be uttered. So, the features of both BIRD and FLYING remain active during the production of "The bird flies." Because of the static message, there is semantic input to both *bird* and *fly* at the same time, that is, there is syntagmatic interference. The model's learning algorithm reduces this interference by developing the syntactic "traffic cop," the configuration of connection weights from syntactic units to lexical items. Figure 3 illustrates the mechanism by presenting some of the weights from the implemented model. Because the message is static, it is up to the syntactic-sequential units to keep track of where the formulation process is in the sentence. In essence, these units direct the traffic, at one point actively encouraging nouns (e.g. the positive weights from the *Noun* unit to all nouns) and inhibiting all non-nouns (e.g. inhibitory weights from *Noun* to verbs and *the*), while at the next point exciting the verbs and inhibiting the other categories. Thus, *bird* is chosen when the current syntactic-sequential state requires the subject noun, and *fly* when it is time for the main verb.

The DoL Model and Semantic/Syntactic Interference

In the DoL model, the syntactic traffic cop and semantic similarity jointly influence the activation of competitors. Recall that in Vigliocco et al. (2005), when pictures were named with inflected verbs, there were two significant main effects on response times: semantically-related distractor words led to longer response times than less related distractors, and verb distractors were more interfering than noun distractors. Although the DoL model is too small and abstract to attempt a quantitative fit of the data, it is demonstrably consistent with these two main effects. Let us use the verb *fly* as an example target in the model, and examine the activations of potential verb and noun competitors. These activations are assumed to be monotonically related to each non-target's propensity to lead to interference in naming time if it were presented as a distractor. To model the naming of the inflected verb condition—the study was done in Italian and the verbs were inflected for third person, present tense—we assumed that there is input from the syntactic-sequential state for the main verb just as if the verb were being produced in a sentence, as well as input from the semantic features of the target. In fact, as Italian is a pro-drop language, the inflected verbs were often full sentences corresponding to, e.g. *It flies*. These are the

activations of the relevant competitors: semantically-related verb, *go* (.508), semantically unrelated verb, *stay* (.305), semantically related noun, *bird* (.059), semantically unrelated noun, *girl* (.038).[2] As long as there is a monotonically increasing relationship between competitor activation and interference in response time, the model produces both the main effect of semantic similarity and the main effect of syntactic category. Thus, the model exhibits the significant effects in the data.

As mentioned before, Vigliocco et al. (2005) also included a condition in which verbs were named in citation (infinitival) form rather than inflected form, and found that the main effect of syntactic category was eliminated. To model this condition, we removed the input from the syntactic-sequential state for the main verb, and again looked at the activation of the four competitors when the target was *fly*: *go* (.226), *stay* (.024), *bird* (.254), *girl* (.167). As in the data, the only sizable pattern is the main effect of semantic similarity. This contrast between the citation and inflected-form simulations illustrates the influence of the syntactic traffic cop. When the cop is on duty, that is, in the inflected condition, it is actively signaling for verbs and inhibiting nouns. When it is not, nouns and verbs compete on an equal footing. We suggest that the difference in the conditions in the data is just the presence versus absence of a verb-signaling cop.

The Syntactic Category Constraint on Aphasic Semantic Errors

The model's syntactic traffic cop also provides a mechanism for the syntactic category effect in speech errors. Each syntactic state tends to excite all words consistent with it and inhibit those that are inconsistent. As a result, words in the wrong category receive little activation and have little chance of being produced. Let us illustrate this by using model lesions to simulate variability in the extent to which semantic substitutions obey the syntactic category constraint, as was found in Berndt et al.'s (1997) study of object and action naming in aphasia. To do this, we describe three additional assumptions in the model. First, we need an explicit mechanism to turn the activation levels of the lexical units into response probabilities. The rule that Gordon and Dell (2003) used was this: activations, a, of each lexical unit are transformed to $e^{\mu a}$, and the probability that each unit is selected is given by the ratio of its transformed value to the sum of the transformed values of all lexical units, with μ (a scaling parameter) = 10. This is a quick-and-dirty method of calculating probabilities that would result from adding noise to the activations, and then picking the most activated unit, with greater values of μ associated with smaller levels of noise. Second, we need to say how object and action naming takes place. Earlier, we suggested that some patients may be instantiating appropriate grammatical contexts when naming nouns or verbs. This is similar to an assumption that Gordon and Dell made when using the model to characterize aphasic single-word production, specifically that the patients are really producing short elliptical sentences, such as (*It's a*) *cat* or (*It's*) *running*. Thus the "message" for a pictured object corresponds to the features of that object and an appropriate syntactic-sequential state, e.g. singular noun. For a pictured action, the message contains the features of an action concept and the syntactic-sequential state is set to that of a verb. Thus, we assume the existence of both semantic and syntactic input to this naming task. Finally, we adopt Gordon and Dell's assumption that patients with agrammatic speech (e.g. Berndt et al.'s FM and LR) may have reduced syntactic input to the lexicon (simulated by low syntactic-lexical weights). Patients who make many lexical errors, but without agrammatic

---

[2] In the implementation, nouns and verbs did not share features. So, a "semantic" relation between a noun and a verb was just the result of their association. For example, *fly* co-occurred with *bird*, but not with *girl*, as the model was trained. These co-occurrences are reflected in the acquired weights and this is why *bird* is more active than *girl* when *fly* is the target. The DoL approach, though, does not in principle prohibit nouns and verbs from sharing features (e.g. both *fly* and *bird* could share the feature, *in_the_air*). If they did, the differences in activation between *bird* and *girl* here would be greater.

speech (e.g. HF and HY), are assumed, instead, to have reduced semantic-to-lexical connections (see also Schwartz et al., 2006).

A normal or unlesioned version of the model follows the syntactic category constraint when naming verbs or nouns in the manner described above. For example, if the target is the verb *fly*, the model is correct 90.3 % of the time, and it generates a semantic error that is also a verb (e.g. *go*) 9.2% of time and a semantically related noun (e.g. *bird*) only 0.12% of the time. The category constraint is even upheld when the semantic-to-lexical weights to zero. In such a model, the proportions of correct responses to the verb *fly*, semantic errors that are verbs, and semantic errors that are nouns are 6.9%, 51.5%, and 0.15%, respectively, in such a model. So, as long as the syntactic-to-lexical connections are intact, the model's errors are constrained to obey the category constraint. In contrast, a model whose syntactic-to-lexical connections are reduced to zero obeys the category constraint to a much lesser extent. Such a model is generally correct when naming the verb *fly* because of its intact semantic-to-lexical connections (89.4% correct), but it makes 1.9% related noun errors (e.g. "bird") to go along with 4.0% related verb errors. Considering all verb and noun errors on the target *fly*, only 61% of the errors matched the verb category.

The model's adherence to the category constraint is due to the syntactic traffic cop, that is, the weights from syntactic-sequential states to lexical items. Aphasic versions of the model that leave these intact generate semantic errors that exhibit the constraint, as did some of Berndt et al.'s patients. Aphasic models that damage these weights create more semantic errors that violate the constraint, as occurred with other patients.

Division of Labor between Syntax and Semantics in Lexical Access

The key property of the DoL model is that its connection weights are learned by attempting to retrieve words incrementally in service of expressing a message. Specifically, the weights are set by the delta rule, a connectionist learning algorithm in which the change in weight from an input to an output unit is proportional to the product of the input's activation and the *error signal* on the output unit. This error signal is itself proportional to the difference between the desired activation of the output and the actual output. So, the learning rule will increase weights from active inputs to underactivated outputs, and decrease weights from active inputs to overactivated outputs.[3]

The delta rule creates *cue competition*, the tendency for inputs to compete to control output. For example, if a model using the delta rule is simulating a conditioning experiment in which cues precede the delivery of food, the rule would determine the weights from each cue (input) to the prediction that food will come (output). If there are two simultaneously available cues that validly signal that food is coming, such as a tone and a light, the rule will cause the cues to compete to predict the food. If one connection weight, say, the one representing prediction of food by the light, happens to become strong, the other will necessarily stay weak. When one input does a good job of predicting the food by itself, the other input's weight cannot grow even though that input, too, is predictive. This is because the delta rule does not allow for additional weight change when there is no longer any error signal.

---

[3] Where does this error signal come from? Production is not usually a supervised task, that is, one in which the correct answer (e.g. no, you should have said "bird" after "the") is supplied by the external environment. Like others (e.g. Chang et al., 2006), we assume that error signals ultimately arise during prediction when doing input processing. When one hears, "The bird flies," in a context in which one can infer the message, a production process attempts to predict each subsequent word. When this prediction conflicts with what is then heard (e.g. the model retrieved "flies" after "the", but it then heard "bird," instead, after "the"), an error signal that can change weights through the delta rule is generated. The process of prediction during input processing transfers completely to production because it is, essentially, production.

In the DoL model, cue competition between semantic and syntactic inputs to activate lexical units leads to differences, or a "division of labor", in the extent to which syntax and semantics participate in the retrieval process. For a word with semantic content such as *bird*, message features such as *animal* are highly predictive cues for that word, and hence semantic-to-lexical weights come to bear much of the responsibility for its access (e.g. a weight of +1.21 from *animal* to *bird*). A function word such as *the* lacks semantic features, however, and so connections from the syntactic-sequential states must pick up the slack. The difference can be seen in the strengths of the syntax-to-lexical connections. There is a weight of +3.44 from the *Det* state to *the*, but only a weight of +1.33 from the *Noun* state to *bird*. In this way, the model has learned how to make up for the semantic impoverishment of *the*, by asking the syntax to do the job. Although the syntactic weight to *bird* is substantial, it does not need to be so large because of the strong semantic input. In this way, the model's learning algorithm leads to differences in the extent to which function and content words are accessed through syntax and semantics, respectively.

The syntactic-semantic competition that leads to differences in the model's representation of function and content words predicts that similar differences will be observed within word categories. Consider the difference between semantically light and heavy verbs. Light verbs such as *go* are assumed to be semantically primitive, perhaps associated with a single feature such as *motion*. A heavier verb such as *fly* can be assumed to have more features, such as *motion* and *through_air* (Breedin, Saffran, & Schwartz, 1998). Along with these complexity differences, light verbs tend to be more frequent and to appear in more contexts (e.g. *boys*, *girls*, *birds*, and *planes* can all *go*, but only *birds* and *planes* can *fly*.) Gordon and Dell (2003) implemented these differences between light and heavy verbs in the model and the result was that, like the function and content words, syntactic and semantic inputs were differentially important for access of the two kinds of verbs. The differences are illustrated in Figure 4, which shows the weights that arose for *go* and *fly*. Clearly, the major source of activation for *go* is the syntactic state for *Verb*, while the major source for *fly* is the semantic feature *through_air*. Notice, in particular, that the syntactic weight for go is more than three times stronger than that for *fly*.

This division of labor is more fully brought out when the model is lesioned. If syntactic and semantic lesions are applied to the model, the differences in the representations of the light and heavy verbs become apparent. Reducing the syntactic-to-lexical weights hurts access of light more than heavy verbs, while reducing semantic-tolexical weights does the reverse. For example, setting the syntactic weights to zero in a simulation of sentence production, leads to 35% correct access for heavy verbs and only 21% correct for light verbs. Reducing semantic weights to zero reverses the effect: 6% correct for heavy verbs and 38% for light verbs (Gordon & Dell, 2003).

Recently, Barde, Schwartz, and Boronat (2006) tested the model's prediction that production deficits associated with syntactic sequencing would impair the production of light verbs more than heavy verbs. They compared two groups of aphasic patients, one labeled "agrammatic" and one "non-agrammatic", in their abilities to retrieve light and heavy verbs in a story telling context. The two groups were matched on their mean semantic comprehension and overall picture-naming abilities. The speech of the agrammatic patients, though, was far less fluent, less syntactically complex, and more lacking in inflectional morphology than that of the other patients. Thus, these patients are assumed, within the context of the DoL model, to have impaired input from syntactic-sequential states to lexical items. (It is not relevant whether the states themselves or their links to the lexicon are impaired). The patients in the study heard short passages including critical sentences such as "Henry *came/drove* home," in which the verb was either light (*came*) or heavy (*drove*). They then were asked, "What did Henry do?" and their spoken responses were scored for the

extent to which they correctly reproduced the verb. As predicted, retrieval success interacted with light/heavy status; the agrammatic group was overall significantly worse on light verbs (31% correct) than heavy verbs (46%). The non-agrammatic group's findings were slightly in the opposite direction (light = 57%; heavy = 54%). More importantly, all patients but one in the agrammatic group (11 of 12) did better on the heavy verbs than on light verbs, often by a large margin, in contrast to the non-agrammatic group, where none of the individuals exhibited much of a preference in either direction. Given that the agrammatic and nongrammatic groups were equally able to understand words, these differences must have arisen post-semantically, that is, in the mapping from meaning to lexical items, which is the locus of the mechanism for the difference proposed by the DoL model. These results are consistent with critical assumptions of the DoL model, specifically the role of learning in setting up semantic and syntactic connections that allow for the access of words in sentence contexts.

A Learning Account of Recurrent Perseverations and Semantic Blocking

Thus far, we have argued that applying models of word production to sentences requires understanding how lexical-access mechanisms deal with syntagmatic as well as paradigmatic interference. We hypothesized that both sources of interference are present at the same production stage and that mechanisms such as the syntactic category constraint help to remove interference from words that do not match the target category. These ideas were made concrete in the DoL model, a network that learned to produce word sequences from a static message and a changing set of syntactic-sequential states. A key feature of this model was the syntactic traffic cop, a learned configuration of excitatory and inhibitory connections from syntactic units to lexical items that implements the syntactic category constraint. Learning in this model also naturally led to a division of labor between semantic and syntactic influences on the retrieval of individual words.

The final section of the article is a further exploration of the forces behind syntagmatic and paradigmatic interference. We examine learning more directly by modeling the changes to the production system that, we propose, result from the production of a lexical item. These changes can cause *recurrent semantic perseverations* in aphasic picture naming. Recurrent perseverations (e.g. Cohen & Dehaene, 1998) are substitutions of a target name with a name that was produced, either correctly or incorrectly, on a previous trial. If the erroneous name is itself semantically related to the target, then the error is a semantic perseveration. Notice that these errors reflect both paradigmatic (semantic) and syntagmatic (previous production) interference.

Although individuals with aphasia can in principle produce semantic perseverations during sentence production, the recurrent perseverations readily induced in picture naming are a more plentiful source of data. We examined 302 semantic errors made by 50 aphasic individuals in a study of picture naming errors by Kittredge, Dell, and Schwartz (2006, submitted), and found that 36% of them were perseverations, that is, a word produced on one of the prior trials. The question then arises as to what causes these errors. It is uncontroversial that the prior production of a word strengthens it in some manner, and that this strengthening makes a perseveratory substitution more likely than a non-perseveratory one, everything else being equal (Martin, Roach, Brecher, & Lowery, 1998; Gotts & Plaut, 2004). But what is the process by which production strengthens a word? We claim that at least part of this strengthening is the result of implicit learning. Insight into this learning comes from research investigating the *semantic blocking effect* in picture naming. When pictures from the same category are repeatedly named by normal speakers (*homogeneous condition*), response times are slowed compared to a condition in which the category varies (*mixed condition*) (e.g. Brown, 1981; Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart,

1994). We focus on a study of this type by Schnur et al. (2006) that examined errors as well as response times.

Schnur et al. (2006)'s Study of Semantic Blocking and Perseverations

Schnur et al. (2006) tested aphasic and control speakers on a variant of the blocking procedure called blocked-cyclic naming, in which a small set of pictures, either a semantically homogeneous or a mixed set, is named in a block and the block is repeated several times. For example, in a homogeneous condition, the speakers might name, "bear," then "raccoon," "lion," "deer," and "goat", and then would get the same items again in a different order, and then again in yet another order, and so on. In the mixed condition, a group of items would be repeated in the same way, except that they would all come from different categories.

The blocking effect in Schnur et al.'s controls showed up in their response times. On the first occurrence of each word in the set (cycle 1), naming response times did not differ between the homogeneous and mixed conditions, but on later cycles, naming in the homogeneous set was slower. For the patients, there were three noteworthy effects and all concerned errors rather than response time. First, a semantic blocking effect emerged in the overall error rates. Analogous to the response-time data with the controls, errors were equally likely in the homogeneous and mixed conditions on the first cycle, but became more likely in the homogeneous condition as more and more cycles were experienced. Second, the difference between the homogeneous and mixed condition error rates was specifically associated with within-set intrusions and omissions. So, repeatedly naming a small set of items from a single category led to relatively more intrusions and omissions compared to repeatedly naming a heterogeneous set of items. Note that in the homogeneous condition, a within-set substitution is necessarily a semantic error and that, aside from the first cycle, such an error is a semantic perseveration. The third important finding concerned these perseverations. Lee, Schnur, and Schwartz (2005) analyzed the semantic perseverations from Schnur et al.'s homogeneous condition and found a lag effect. For example, perseverations were more likely to be substitutions of responses that occurred, say, 2 trials ago than 6 trials ago. But this lag effect, surprisingly, is not due to temporal decay of the potential for a response to perseverate. In the Schnur et al. study, naming trials either occurred with a 1-second or 5-second response-to-next-stimulus (RSI) interval. If the lag effect were due to time, the lag functions should be much steeper when more time occurs between trials. Instead, the lag functions and, more generally, the overall rate of semantic perseverations were unaffected by RSI. Intervening trials, not the passage of time, are what matters. These findings suggest that, at least in this experiment, the relevant changes to the lexical system that were responsible for interference from previous items and the resulting perseverations are the product of the implicit learning that occurs during the production of the previous words. That is, this change is relatively more permanent, and not a time-decaying activational process (see Bock & Griffin, 2000 for further discussion of the distinction between learning and activation).

A Model of Schnur et al.'s (2006) Data

The perseveration model (Figure 5) uses the error-based learning approach of the DoL model to account for the qualitative properties of Schnur et al.'s findings. Like the DoL model, there were semantic features connecting directly to words. However, we did not implement syntactic input because all targets and sources of interference were nouns. We first trained the network to produce its targets correctly, by training each target many times using the delta rule. Figure 6 shows the learned connection weights for example inputs after training. Then we simulated the blocking manipulations. Homogeneous conditions were simulated by repeatedly accessing a block of 3, e.g., mammals (which shared the feature

mammal), and mixed conditions by repeatedly accessing a block of 3 unrelated targets (sharing no features). As in training, the model's connection weights continued to be modified during these simulations.

As with the application of the DoL model to aphasic errors, we need a method to turn activations into responses. For the DoL model, we used a mathematical shortcut that simulates a decision process in which noise is added to each potential output and the most activated output is chosen. For the perseveration model, we also added noise to the activations and picked the winner. But we did so directly because we also needed to specify a mechanism for omission errors and response times, in order to attempt to simulate the properties of Schnur et al.'s data. Specifically, the noisy word activation values were turned into responses by the following rule: if the most activated word is greater than the mean of the others by difference threshold of .99, that word is produced. If it is not, then a *booster* mechanism kicks in. The booster, which is assumed to represent either a general or linguistically-specialized mechanism for resolving competition (Thompson-Schill, D'Esposito, & Kan, 1999), repeatedly multiples all word activations by a boosting factor (1.01) until the difference threshold is reached. The number of multiplications needed adds to the response time, and an omission is assumed to occur if the booster fails to make a word reach the threshold after 100 multiplications. The values of the difference threshold, the boosting factor, and the omission time-out permit considerable variation. The model will exhibit the qualitative features that we describe below provided that the booster is greater than 1, but sufficiently close to it that it takes a variable number of boosts to hit the difference threshold. It must also be assumed that a fair number of trials fail to reach the threshold before the omission time-out is reached.

Simulation Results

The model accounts for the general findings of Schnur et al. The difference between aphasic and normal speakers was simulated by adding extra noise to the activation levels for the aphasic condition. The normal simulations exhibited a semantic blocking effect in the response times that increased over cycles (Figure 7). This arises because the model contains the three properties that Howard, Nickels, Coltheart, and Cole-Virtue (2006) claimed were needed to produce this response-time effect: a mechanism by which previous responses are strengthened (weights from features to targets are strengthened and weights from features to activated nontargets are weakened on each trial in our model), a mechanism that leads to the activation of semantically-related words as well as the target word (the activated feature *mammal* sends activation to nontarget mammal names as well as the target), and a mechanism that leads to longer response times when nontargets are more activated (the difference threshold for responses). In short, the previous production of *dog* made *dog* more accessible and *bat* less accessible from *mammal*. So, when *bat* is now the target, the activated shared feature *mammal* sends more activation to *dog* than to *bat*, and hence the response time for *bat* is longer than it would normally be.

In the aphasic simulation, the homogeneous condition is associated with more semantic errors (largely perseverations) and omissions than the mixed condition for much the same reasons (Figures 8 and 9). Strengthening *dog* increases the chance that it will substitute for *bat*, and even if it does not substitute for *bat*, it increases the chance that the booster fails to boost *bat* over the difference threshold and, hence, that an omission occurs. Finally, the model produces a perseveratory lag function that is, as in the data, sensitive not to RSI, but only to the number of intervening semantically related trials. After *dog* is produced, the degree to which it is strengthened is diminished by the subsequent production of *bat*, due to the error-driven nature of the learning algorithm. Specifically, on the production of *bat*, *dog* is erroneously activated to an extra extent, and hence, the weight from *mammal* to *dog* is diminished to a greater extent. So, when *whale* is next produced, *bat* is a more potent

competitor than *dog*. These perseverative influences in the model are, like the data, uninfluenced by the passage of time. In addition, the influences of the production of one member of a semantic category on a member of the same category are, in the model, not affected by intervening semantically *unrelated* trials because, on such trials, no weight change occurs on connections involving the shared features of the category. This is exactly what Howard et al. (2006) found using a blocking paradigm in which a variable number of unrelated trials intervened between repetitions of a category. Our model is insensitive to time and unrelated trials because the perseverations are due to weight change, that is, to learning. Like Howard et al., we suggest that the perseverative impetus in this paradigm is not the result of a rapidly decaying change such as would be expected with an activation or working-memory account of perseveration.

In summary, the paradigmatic and syntagmatic interference that leads to semantic perseverations can be understood by assuming that both sources of interference impact the same component of lexical retrieval and that each lexical retrieval act involves learning. Just as in the DoL model, we claim that at least some component of that learning involves error correction. For the DoL model, the error correction helped to explain the differences between light and heavy verbs. For the model of the semantic blocking effect, error correction explained the effect of lag on semantic perseverations, and the absence of an effect of RSI on these errors.

Conclusions

The theme of this article is that the lexical access system learns to overcome the interference that it experiences when producing words in context. It must deal with paradigmatic interference from semantic competitors, and syntagmatic interference from previously produced and upcoming words. This learning leads to particular mechanisms, among them the syntactic traffic cop, which enforces the syntactic category constraint. Furthermore, the models that we discussed use a particular kind of learning, error-correction learning, and this enables them to explain some aspects of word and sentence-production data. Thus, learning plays an important role in accounting for both the effects of syntactic and paradigmatic interference and the mechanisms developed by the production system to resolve this competition (e.g. syntactic traffic cop).

Admittedly, these models are overly simple. They leave out much of the complexity that makes production so interesting. And there is no true sense in which either of the models stands as an account of language acquisition. Their use of learning is limited to weight adjustments in specific sets of connections among semantic, lexical, and syntactic representations that are already in place. Our goal has been to understand the role of learning in explaining mechanisms of lexical selection, rather than to account for developmental data. It is possible, however, to develop larger-scale learning-based models, models that confront more of the complexities of production and which better respect the challenges facing theories of acquisition. In fact, the Dual-path recurrent network model of structural priming by Chang et al. (2006; Chang, 2002) embodies, at a greater level of complexity, the major assumptions of the DoL model. Both models separate lexical-semantic representational pathways from syntactic-sequential states, and both use error-driven learning to control how these two information sources coordinate incremental lexical access from a static message. As a result, the Dual-path model implements its own learned syntactic traffic cop in its connections from the sequencing system to lexical output. For example, the activation patterns in the hidden units of the model's sequencing system correspond to syntactic categories and subcategories that are more finely tuned to sentence structure than just the global *Noun* or *Verb* states assumed for the DoL model (Chang, 2002). It's as if the cop has many arms to control a multi-way intersection. Unlike the simple DoL model, though, the

Dual-path model's syntactic-sequential states and how they change, as well as their connections to lexical items, are learned by the error-based learning algorithm.

The Dual-path model also differs from our models by covering more complex sentences and production phenomena, including subject-verb agreement, a wide variety of sentence and noun-phrase structures, and the influence of meaning on structural selection. As with our simple models, the learning aspect of the Dual-path model was tested in two ways, by the nature of the representations it develops through learning (e.g. the division of labor between the lexical semantic and sequential systems), and through the implicit learning phenomena that it explains (structural priming in the case of the Dual-path model; lexical perseverations in our perseveration model). In addition, the Dual-path model directly simulates language acquisition data, such as differences in when children acquire particular constructions. Moreover, like the child, it learns from processing input and then transferring that knowledge to the task of production.

Although the Dual-path model has many features in common with the models presented here and is clearly superior in many respects, it has limitations that prevent it from being the focus of our present analysis. Unlike the DoL model, it has not previously been applied to error data and, even if it were so applied, it lacks semantic representations that can be the basis for semantic similarity effects (e.g. shared features). So, the data concerning speech errors and particularly semantically based errors and response time effects are outside of its domain. Given its complexity, it is also quite difficult to understand the Dual-path model's mechanisms, unlike the models that we have focused on here, whose simplicity makes their mechanisms transparent. Each of the mechanisms that we have identified, such as the traffic cop, the division of labor between syntax and semantics, and the perseverative influences resulting from multiple retrievals from the same semantic category, are clearly visible in the learned excitatory and inhibitory weights. In this way, the simple and complex models complement one another. The simple model clarifies theory, while the complex one speaks to scalability and coverage. We feel that progress in extending word-production models to sentence production will be promoted by models at various levels of abstraction.

Ultimately, a full understanding of sentence production will require a greater consideration of processes that are not the traditional domain of production research. These include the processes that focus attention on the world and our conceptions of it (e.g. Bock, Irwin, Davidson, & Levelt, 2003), that control selection under competition (e.g. Roelofs, 2003), that guide the interplay of perception and action (e.g. Hartsuiker & Kolk, 2001; Slevc & Ferreira, 2006), and, as we have emphasized here, the processes by which we acquire and hone our production abilities. In short, it requires that production be placed within the matrix of human cognition.

Acknowledgments

References

Alario F-X, Caramazza A. The production of determiners: Evidence from French. Cognition 2002;82:179–223. [PubMed: 11747862]

Barde LH, Schwartz MF, Boronat CB. Semantic weight and verb retrieval. Brain and Language 2006;97:266–278. [PubMed: 16359726]

Belke E, Meyer AS, Damian MF. Refractory effects in picture naming as assessed in a semantic blocking paradigm. The Quarterly Journal of Experimental Psychology: Section A 2005;58:667–692.

Berndt RS, Mitchum CC, Haendiges AH, Sandson J. Verb retrieval in aphasia. Brain and Language 1997;56:68–106. [PubMed: 8994699]

Bock K, Irwin DE, Davidson DJ, Levelt WJM. Minding the clock. Journal of Memory and Language 2003;48:653–685.

Breedin SD, Saffran EM, Schwartz MF. Semantic factors in verb retrieval: An effect of complexity. Brain and Language 1998;63:1–35. [PubMed: 9642018]

Brown AS. Inhibition in cued retrieval. Journal of Experimental Psychology: Human Learning and Memory 1981;7(3):204–215.

Caramazza A. How many levels of processing are there in lexical access? Cognitive Neuropsychology 1997;14:177–208.

Chang F. Symbolically speaking: A connectionist model of sentence production. Cognitive Science 2002;26:609–651.

Chang F, Dell GS, Bock K. Becoming syntactic. Psychological Review 2006;113:234–272. [PubMed: 16637761]

Cohen L, Dehaene S. Competitions between past and present: Assessment and interpretation of verbal perseverations. Brain 1998;121:1641–1659. [PubMed: 9762954]

Damian MF, Vigliocco G, Levelt WJM. Effects of semantic context in the naming of pictures and words. Cognition 2001;81(3):B77–B86. [PubMed: 11483172]

Dell GS. A spreading activation theory of retrieval in language production. Psychological Review 1986;93:283–321. [PubMed: 3749399]

Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. Psychological Review 1997;104:801–838. [PubMed: 9337631]

Eberhard KM, Cutting JC, Bock K. Making syntax of sense: Number agreement. Psychological Review 2005;112:531–559. [PubMed: 16060750]

Ferreira VS. The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. Journal of Memory and Language 2003;48:379–398.

Fromkin VA. The non-anomalous nature of anomalous utterances. Language 1971;47:27–52.

Garrett, MF. The analysis of sentence production. In: Bower, GH., editor. The psychology of learning and motivation. San Diego: Academic Press; 1975. p. 133-175.

Goldrick M. Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. Language and Cognitive Processes 2006;21:817–855.

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.

Gotts SJ, Plaut DC. Connectionist approaches to understanding aphasic perseveration. Seminars in Speech and Language 2004;25:323–334. [PubMed: 15599822]

Griffin ZM, Bock JK. What the eyes say about speaking. Psychological Science 2000;11:274–279. [PubMed: 11273384]

Hartsuiker RJ, Kolk HHJ. Error monitoring in speech production: A computational test of the perceptual loop theory. Cognitive Psychology 2001;42:113–157. [PubMed: 11259106]

Howard D, Nickels L, Coltheart M, Cole-Virtue J. Cumulative semantic inhibition in picture naming: Experimental and computational studies. Cognition 2006;100:464–482. [PubMed: 16413014]

Kittredge AK, Dell GS, Schwartz MF. Aphasic picture-naming errors reveal the influence of lexical variables on production stages. Brain and Language 2006;99:216–217.

Kittredge AK, Dell GS, Schwartz MF. Where is the frequency effect in word production? Insights from aphasic picture naming errors. (submitted).

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(2):149–174.

Lecours AR, Lhermitte F. Phonemic paraphasias: Linguistic structures and tentative hypotheses. Cortex 1969;5:193–228. [PubMed: 5391877]

Lee, E.; Schnur, T.; Schwartz, MF. Recency of production influences semantic substitutions in blocked-cycle naming. 46th Annual Meeting of the Psychonomic Society; November 10–13 2005; Toronto. 2005.

Levelt, WJM. Speaking: From intention to articulation. Cambridge, MA: MIT Press; 1989.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behavioral and Brain Science 1999;22(1) 991-975.

MacKay DG. The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behaviors. Psychological Review 1982;89:483–506.

Martin N, Roach A, Brecher A, Lowery J. Lexical retrieval mechanisms underlying whole-word perseveration errors in anomic aphasia. Aphasiology 1998;12:319–333.

Martin RC. Language processing: Functional organization and neuroanatomical basis. Annual Review of Psychology 2003;54:55–89.

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;33:503–535.

Nickels, L. Word production. In: Rapp, B., editor. The handbook of neuropsychology: What deficits reveal about the human mind. Philadelphia: Psychology Press; 2001.

Nooteboom, SG. The tongue slips into patterns. In: Sciarone, AG.; van Essen, AJ.; Van Raad, AA., editors. Leyden studies in linguistics and phonetics. The Hague: Mouton; 1969. p. 114-132.

Pechmann T, Zerbst D. The activation of word class information during speech production. Journal of Experimental Psychology: Learning, Memory, & Cognition 2002;28:233–243.

Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. Psychological Review 2000;107:460–499. [PubMed: 10941277]

Roelofs A. The WEAVER model of word-form encoding in speech production. Cognition 1997;64:249–284. [PubMed: 9426503]

Roelofs A. Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. Psychological Review 2003;110:88–125. [PubMed: 12529058]

Ruml W, Caramazza A, Capasso R, Miceli G. Interactivity and continuity in normal and aphasic language production. Cognition Neuropsychology 2005;22:131–168.

Saffran EM, Berndt RS, Schwartz MF. The quantitative analysis of agrammatic production: Procedure and data. Brain and Language 1989;37:440–479. [PubMed: 2804622]

Schnur TT, Schwartz MF, Brecher A, Hodgson C. Semantic interference during blocked-cyclic naming: Evidence from aphasia. Journal of Memory and Language 2006;54:199–227.

Schriefers H, Meyer AS, Levelt WJM. Exploring the time-course of lexical access in production: Picture-word interference studies. Journal of Memory and Language 1990;29:86–102.

Schwartz MF, Dell GS, Martin N, Gahl S, Sobel P. A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. Journal of Memory and Language 2006;54:228–264.

Slevc LR, Ferreira VS. Halting in single-word production: A test of the perceptual loop theory of speech monitoring. Journal of Memory and Language 2006;54:515–540. [PubMed: 17917694]

Thompson-Schill SL, D1Esposito M, Kan IP. Effects of repetition and competition on prefrontal activity during word generation. Neuron 1999;23:513–522. [PubMed: 10433263]

Vigliocco G, Vinson DP, Siri S. Semantic similarity and grammatical class in naming actions. Cognition 2005;94:B91–B100. [PubMed: 15617670]

Vigliocco G, Vinson DP, Lewis W, Garrett MF. Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. Cognitive Psychology 2004;48:422–488. [PubMed: 15099798]

Widrow, B.; Hoff, ME. Adaptive switching circuits. In: Anderson, JA.; Rosenfeld, E., editors. Neurocomputing: Foundation of research. Cambridge, MA: MIT Press; 1988. p. 126-134.
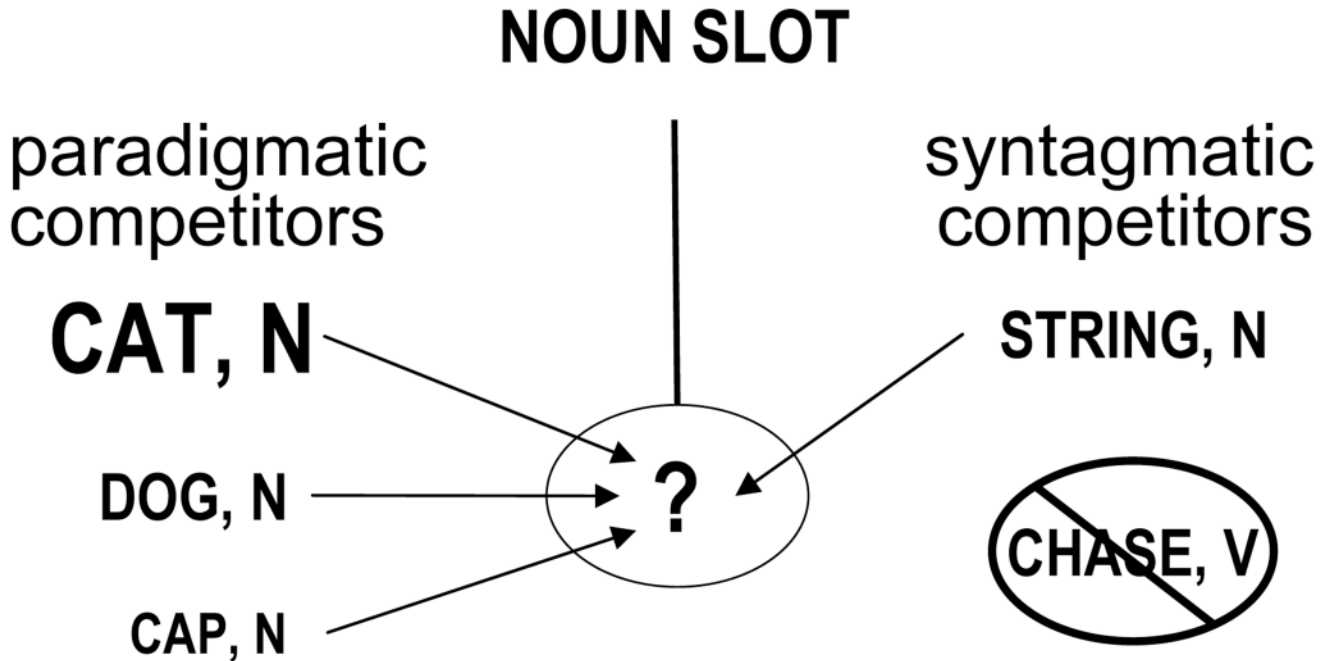
"The **cat** might chase the string"

NOUN SLOT

paradigmatic competitors

syntagmatic competitors

CAT, N

STRING, N

DOG, N

?

CAP, N

CHASE, V

Fig 1.
Paradigmatic and syntagmatic competitors all compete with the target, based on activation levels, and filtered by category.
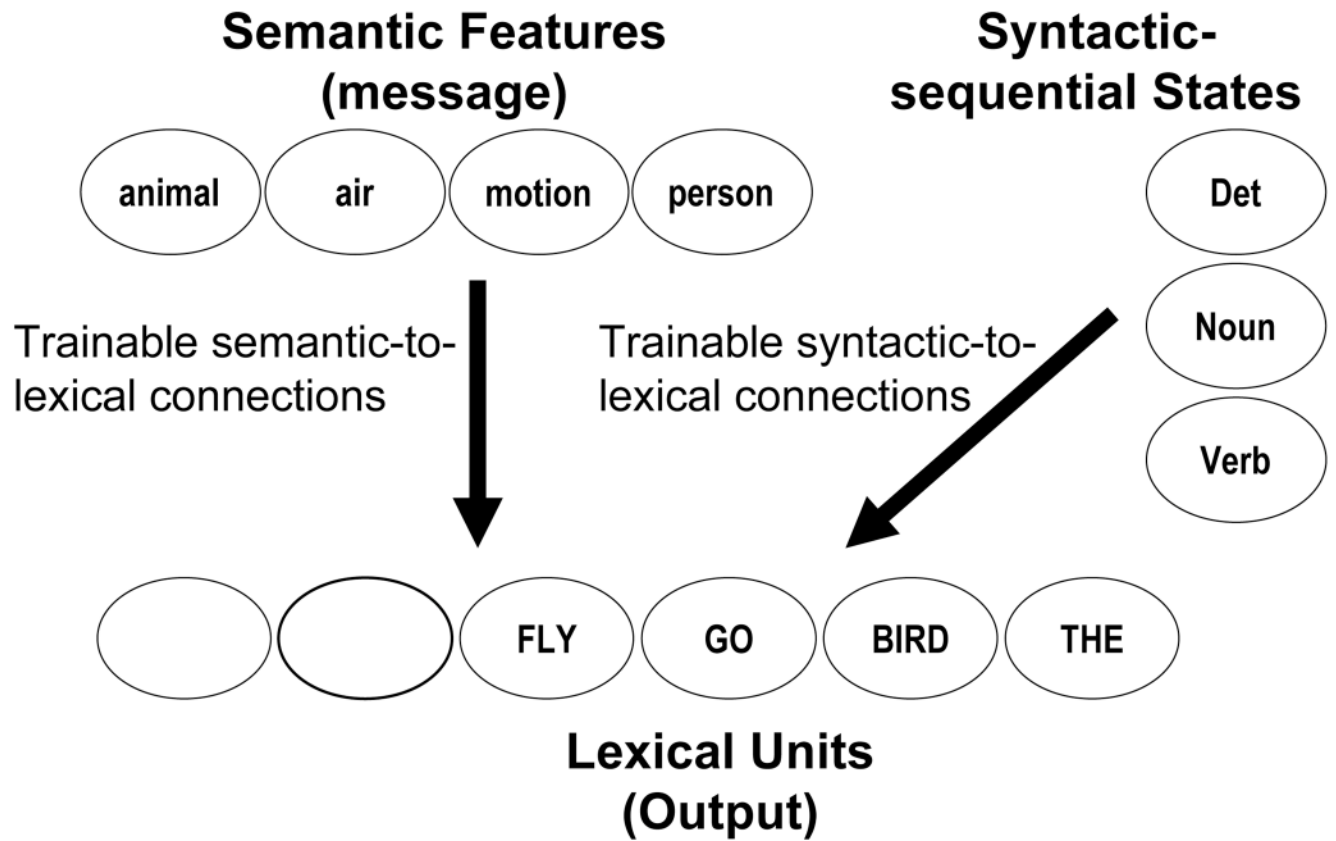
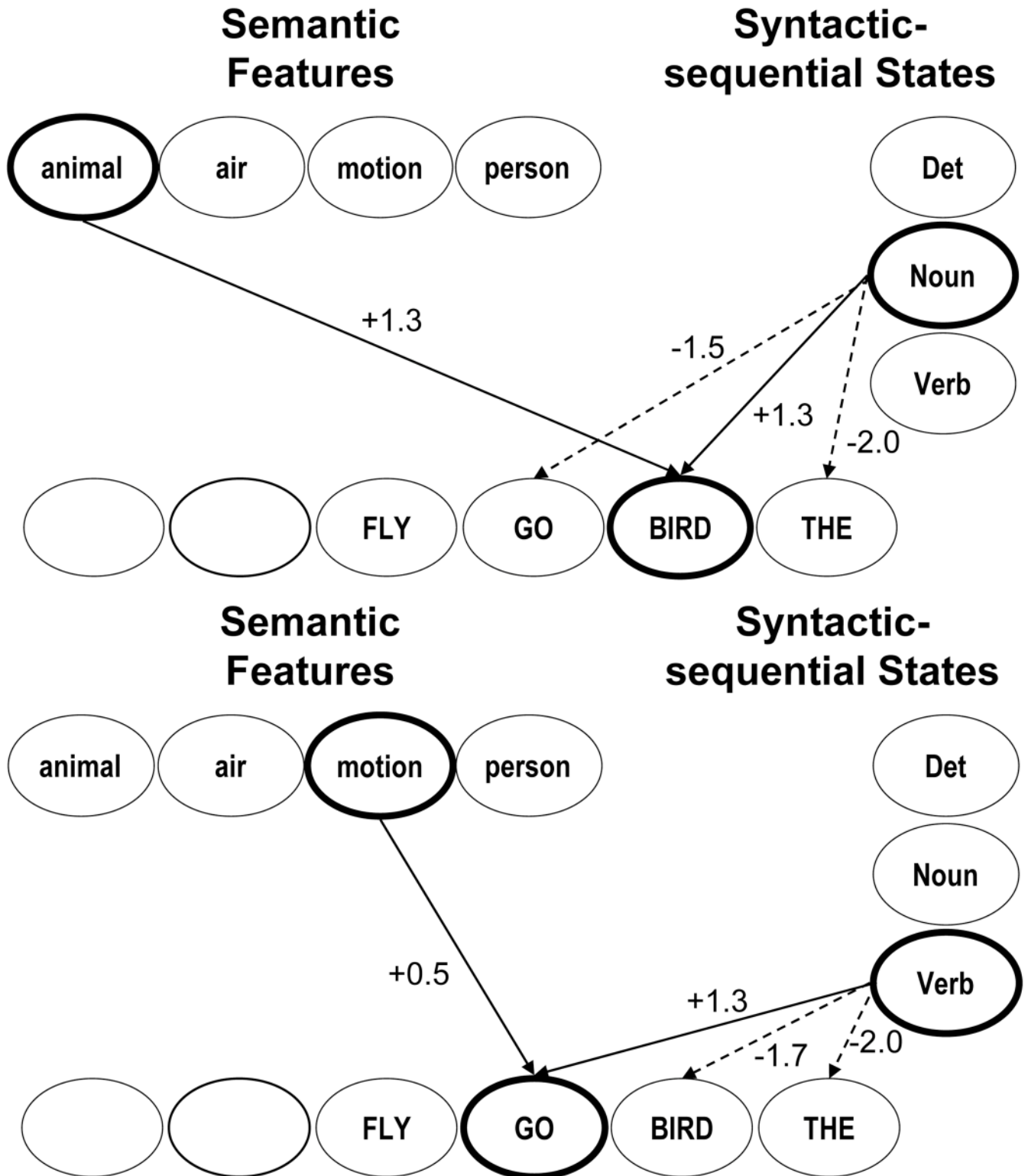**Fig 2.**
The division of labor model.

Fig 3.
Fig 3a. The syntactic traffic cop signaling for nouns.

Fig 3b. The syntactic traffic cop signaling for verbs.
NOTE: Putting Fig's 3a and 3b right next to each other is what the authors intend
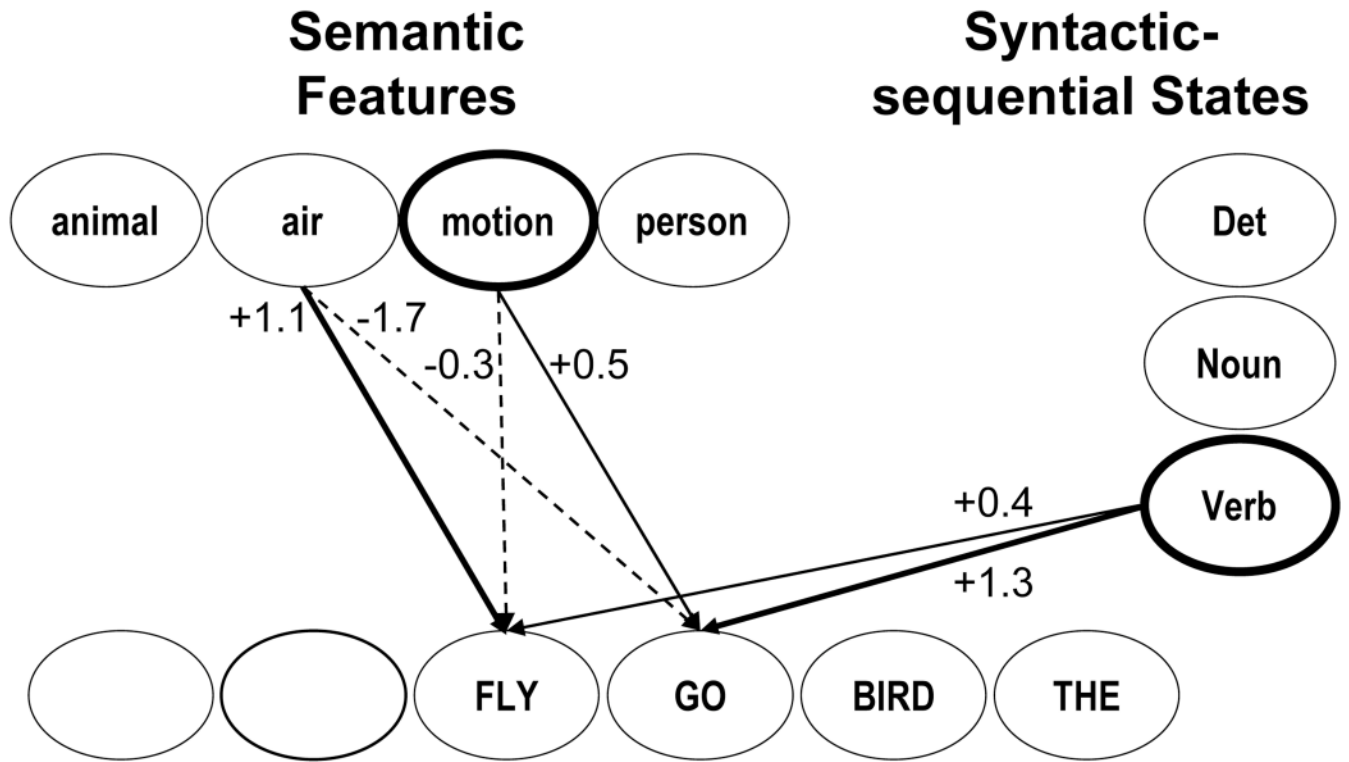
Fig 4.
Light and heavy verbs differentially depend on semantic and syntactic input.
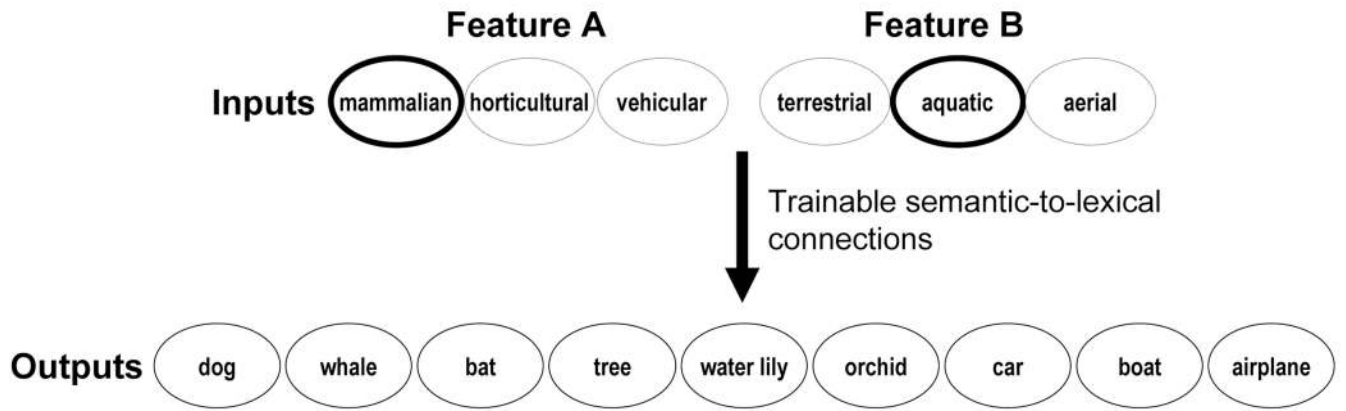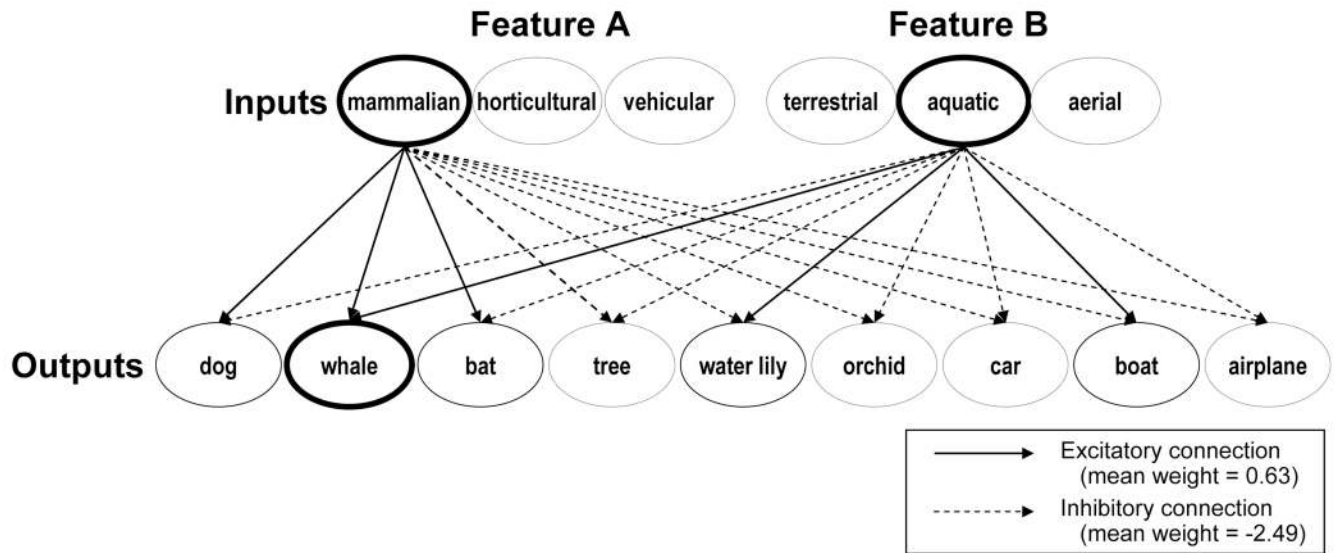
Fig 5.
Architecture of the perseveration model.

Fig 6.
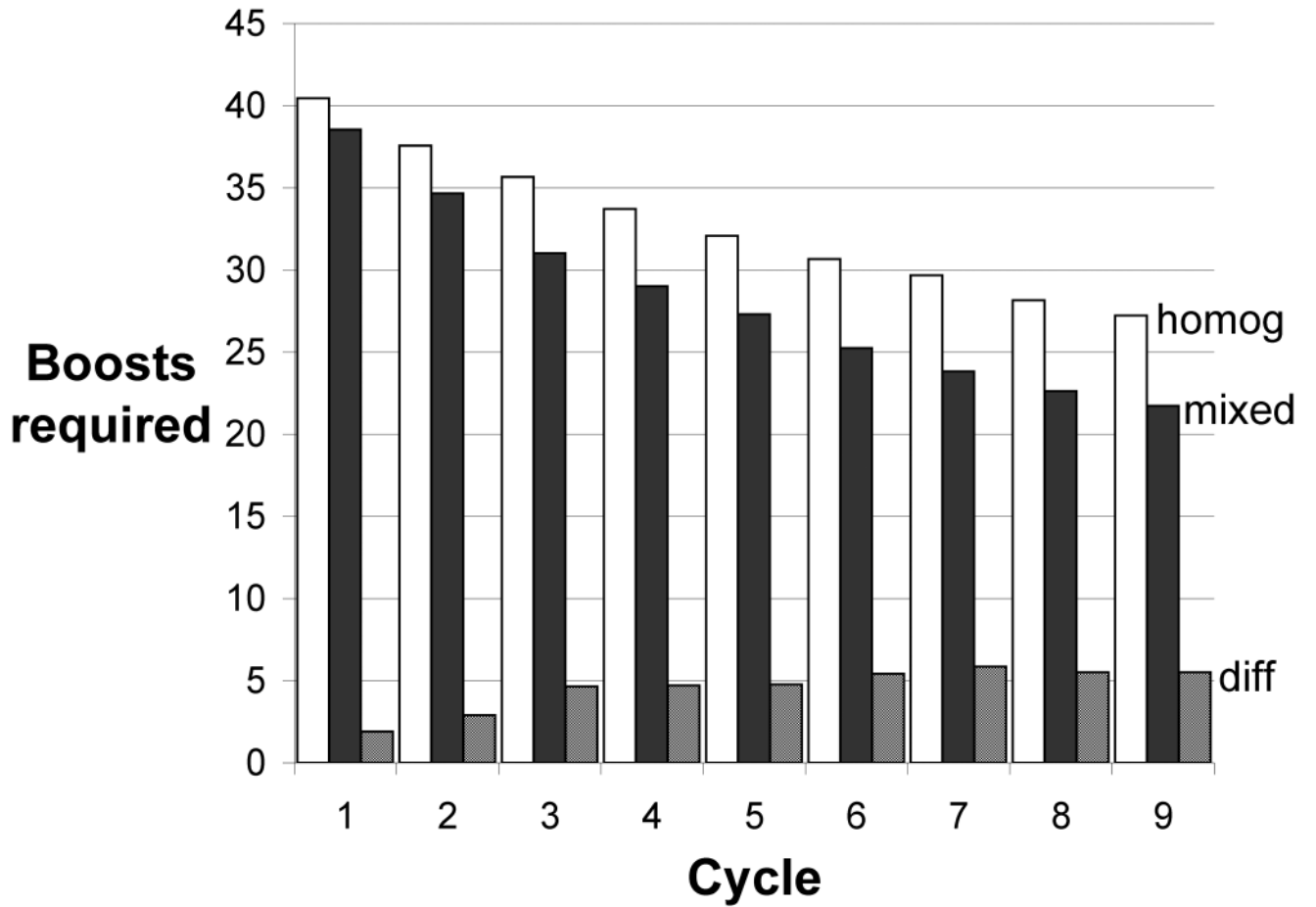Learned activation patterns for an example pair of input features.

Fig 7.
Simulated response times in the perseveration model (number of "boosts" required to reach threshold) as a function of blocking condition; homog = semantically homogeneous block; mixed = mixed block; diff = homog - mixed.
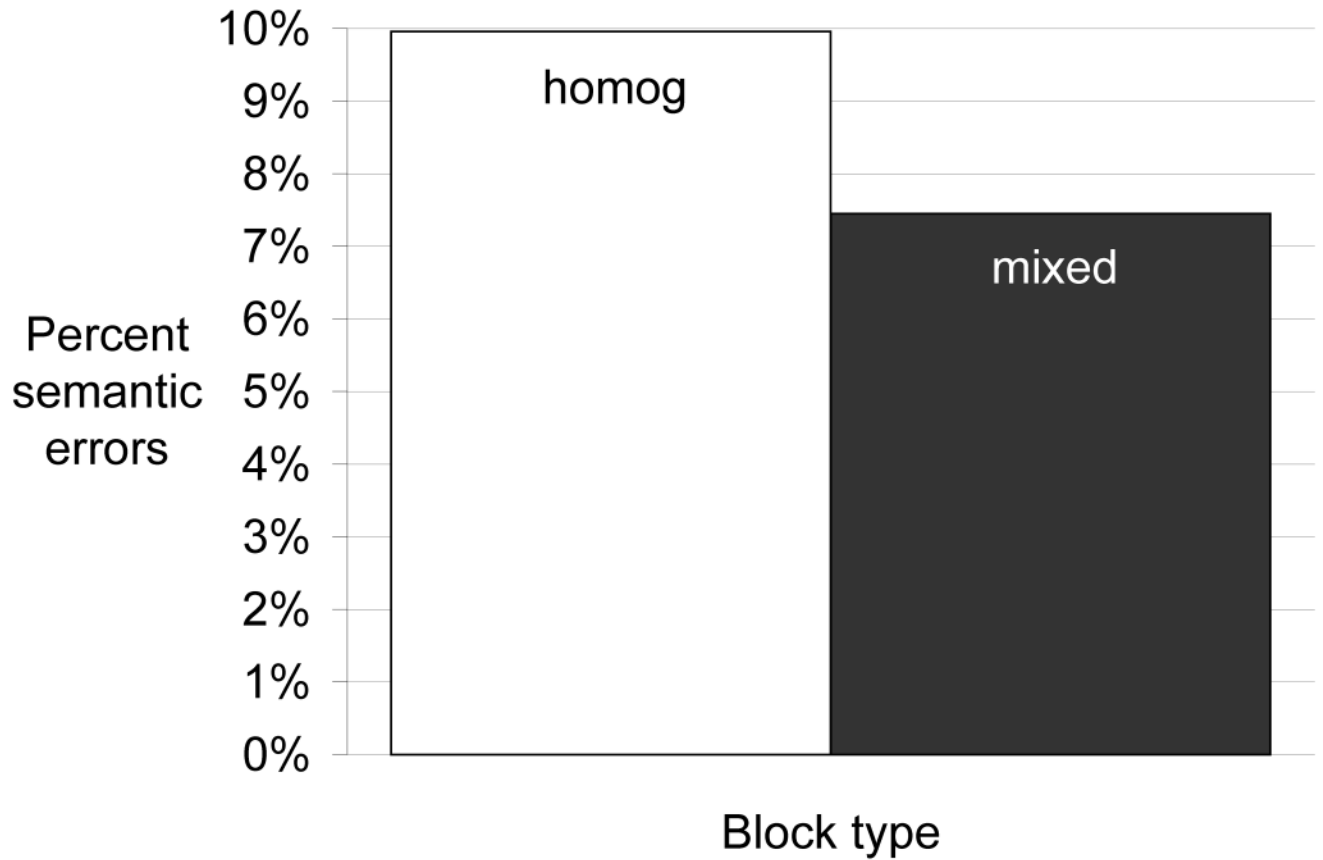
Fig 8.
Percentage of semantic errors in the homogeneous (homog) and mixed (mixed) blocks simulated by the perseveration model.
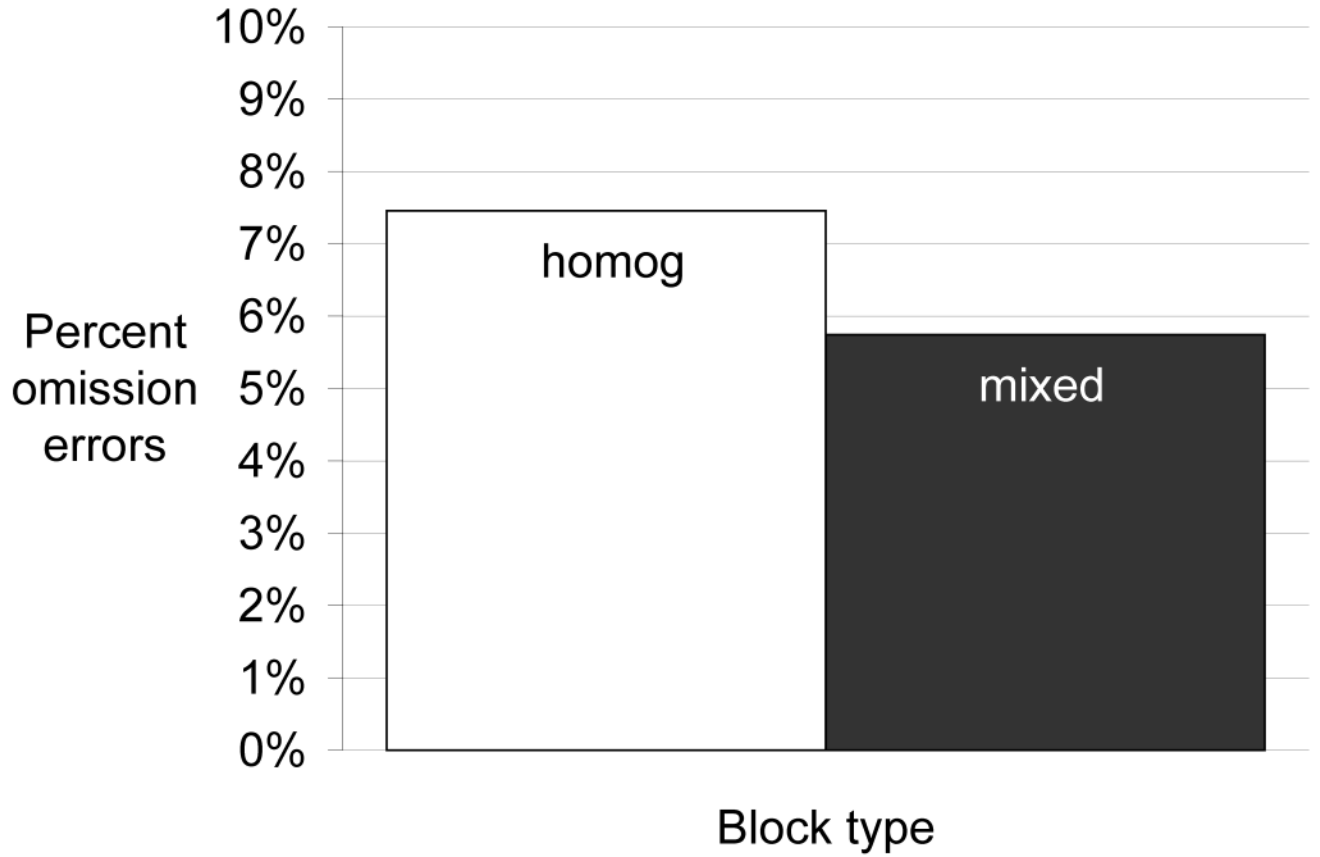
Fig 9.
Percentage of omission errors in the homogeneous (homog) and mixed (mixed) blocks simulated by the perseveration model.