ELSEVIER

Cognitive Science 27 (2003) 1–40

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Learning to divide the labor: an account of deficits in light and heavy verb production

Jean K. Gordon[a,*], Gary S. Dell[b,1]

[a] *Department of Speech Pathology and Audiology, University of Iowa, 125B WJSHC, 250 Hawkins Drive, Iowa City, IA 52242, USA*

[b] *University of Illinois at Urbana-Champaign, Beckman Institute, 405 North Mathews Avenue, Urbana, IL 61801, USA*

Received 2 April 2002; received in revised form 26 July 2002; accepted 31 July 2002

Abstract

Theories of sentence production that involve a convergence of activation from conceptual-semantic and syntactic-sequential units inspired a connectionist model that was trained to produce simple sentences. The model used a learning algorithm that resulted in a sharing of responsibility (or "division of labor") between syntactic and semantic inputs for lexical activation according to their predictive power. Semantically rich, or "heavy", verbs in the model came to rely on semantic cues more than on syntactic cues, whereas semantically impoverished, or "light", verbs relied more on syntactic cues. When the syntactic and semantic inputs were lesioned, the model exhibited patterns of production characteristic of agrammatic and anomic aphasic patients, respectively. Anomic models tended to lose the ability to retrieve heavy verbs, whereas agrammatic models were more impaired in retrieving light verbs. These results obtained in both sentence production and single-word naming simulations. Moreover, simulated agrammatic lexical retrieval was more impaired overall in sentences than in single-word tasks, in agreement with the literature. The results provide a demonstration of the division-of-labor principle, as well as general support for the claim that connectionist learning principles can contribute to the understanding of non-transparent neuropsychological dissociations.

Keywords: Agrammatism and anomia; Aphasia; Connectionist models; Language production; Verb deficits

* Corresponding author. Tel.: +1-319-335-8729; fax: +1-319-335-8851.
E-mail address: jean-k-gordon@uiowa.edu (J.K. Gordon), gdell@s.psych.uiuc.edu (G.S. Dell).
[1] Tel.: +1-217-244-1294.

PII: S0364-0213(02)00111-8

1. Introduction

The production of an utterance is a complex process, involving the convergence of multiple sources of information throughout a succession of stages (Bock, 1982; Dell, 1986; Garrett, 1975; Levelt, 1989). Production begins with a *message*, a meaning to be conveyed. The message guides the retrieval of lexical items and the construction of an appropriate schema or frame that specifies in what position and in what form the lexical items appear. Consider the utterance "The bird flies." The message contains the concepts of a particular BIRD and FLYING, indicating that a previously discussed bird is presently flying. The appropriate lexical items "bird" and "fly" are retrieved and a simple intransitive frame is prepared. This frame indicates that the subject noun is to be placed after the determiner, in accordance with the grammar of English noun phrases, and the main verb, suitably inflected for tense and number, is to follow.

On this view, producing a particular word at a particular point in a sentence requires the integration of semantic and syntactic information. We say "bird" because its semantics correspond to the entity in question, and we say it after "the" but before "flies", because of the syntax of English. That lexical retrieval in production is subject to both semantic and syntactic constraints leads to the hypotheses explored in this article: the relative strength of the respective contributions of semantics and syntax during lexical access varies among words, and this variation can explain certain dissociations in aphasic language production. Most importantly, the syntactic-semantic variation arises from *cue competition* in learning, resulting in a *division of labor* between these two information sources.

Cue competition is the tendency for inputs to compete with each other to control output; the strength of each input cue varies as a function of the number and strength of other inputs involved in determining the output. This can arise from, among other mechanisms, associative learning algorithms that set the weights from input to output so as to reduce error (Rescorla & Wagner, 1972; Widrow & Hoff, 1988). Recent connectionist models have used cue competition to account for differences in the processing of regular and irregular linguistic mappings. For example, Plaut and colleagues (Plaut, McClelland, Seidenberg, & Patterson, 1996) hypothesized that regularly spelled words are read aloud largely on the basis of phonological input, but irregular words require more semantic input. Their model's architecture created competition between phonological and semantic cues to predict orthographic output, and the model's error-driven learning algorithm led to connection weights that "divided the labor" between the two sets of cues in the hypothesized manner.

We apply this division-of-labor logic to the relative contributions of semantics and syntax to lexical access in speech production. Our modeling studies seek to explain neuropsychological data from the study of aphasia suggesting that lexical items are differentially dependent on semantic and syntactic input. Importantly, we do not stipulate *a priori* that certain words rely more or less on semantic or syntactic information. Rather, the differences emerge from the natural constraints on the task of producing grammatical sentences, coupled with the use of a learning algorithm that creates cue competition. When lesioned, the models illustrate the contrast between lexical retrieval deficits typical of agrammatic patients and those typical of anomic patients. Dissociations found in aphasic speech production, particularly between semantically "heavy" verbs (e.g., "fly") and semantically "light" verbs (e.g., "go") (Breedin, Saffran, & Schwartz, 1998), are replicated in the models. The results of the simulations demonstrate that connectionist models can explain double dissociations of symptoms without the need to postulate autonomous representational stores.

The present study links three domains in cognitive science: language production, aphasia, and connectionist models of learning. The principles of connectionist modeling are applied to what we know about normal language production and what we have observed in aphasic patients, in order to clarify the underlying nature of lexical access deficits in aphasia. The following sections review the relevant literature in each domain.

1.1. Lexical retrieval in language production

Production theorists have long acknowledged the impact of syntactic as well as semantic knowledge on lexical access. Syntactic effects are particularly apparent in two properties of everyday speech errors, the syntactic category constraint and the differences between errors involving open- and closed-class lexical items. In word substitution errors, such as saying "knee" for "elbow" or "stop" for "start," or in word exchange errors (e.g., "I'm writing a *mother* to my *letter*"), the interacting words almost always belong to the same syntactic category (Fromkin, 1971; Garrett, 1975). Nouns replace nouns, verbs replace verbs, and so on. A standard interpretation of the syntactic category constraint is that sentence planning entails the insertion of retrieved lexical items into slots in a frame that corresponds to the syntactic structure of the sentence. The slots are specified by their syntactic category with the result that only lexical items of the proper category can be inserted (Bock & Levelt, 1994; Garrett, 1975).

This slot-filler mechanism is also called upon to explain differences in the constraints operating on open- and closed-class words. Closed-class categories—such as prepositions, auxiliaries, pronouns, determiners, and conjunctions—are so called because new words cannot be added to them. In contrast, open categories—such as nouns, verbs, and adjectives—can be expanded, because speakers continue to learn open-class words throughout their lives and have the ability to create new category members through the productive use of morphology. These differences are reflected in speech-error patterns. Garrett (1975, 1988) noted that closed-class words are not involved in exchange errors the way that open-class words are. For example, exchanges such as "I'm writing a *mother* to my *letter*" are common, but "I'm writing *my* letter to *a* mother" is not. For this and other reasons, Garrett proposed that closed-class words are not inserted into slots in the way that open-class words are; rather, they are an inherent part of the frame itself. By locating closed-class words in the frame, Garrett's proposal links their retrieval to syntactic processing, in contrast to open-class words, which are more dependent on semantic information. As such, closed-class words have also come to be known as "function" words, because they serve a predominantly grammatical function and are relatively devoid of meaning. Open-class words are referred to as "content" words, because they convey most of the semantic content of the message (Gleason, 1961). As will be discussed later, this view has been very influential in characterizations of aphasic deficits.

In recent theories of sentence production, the notion of lexical items filling slots in syntactically specified frames has been retained, and the processes have been quantitatively specified in terms of spreading activation. Words are retrieved by the spread of activation from semantic or conceptual units to candidate lexical units in a lexical network (Caramazza, 1997; Cutting & Ferreira, 1999; Dell, 1986; Dell, Burger, & Svec, 1997; Dell, Schwartz, Martin, Saffran,

& Gagnon, 1997; Harley, 1984; Levelt, Roelofs, & Meyer, 1999; Rapp & Goldrick, 2000; Roelofs, 1992; Stemberger, 1985). For example, activation from a message representing the concept BIRD would spread to the lexical unit "bird" and, to a lesser extent, to related lexical units, such as "plane" and "fly". The insertion of "bird" into its grammatically appropriate slot in the developing utterance can also be achieved by spreading activation, provided that the lexical network is augmented by a separate network that can represent syntactic properties, or what we will call *syntactic-sequential states*. These states have two important properties. First, they change throughout the encoding of the sentence, starting at a state representing the beginning of the sentence, moving through structurally defined states, and finishing at an end-of-sentence state (e.g., for the simple sentence "The bird flies", the sequential states would proceed from the noun phrase—in this case, determiner then noun—to the verb phrase). Second, these states send activation to lexical units with which they are consistent (e.g., from the NOUN state to "bird" and to other nouns, like "plane", but not to the verb "fly"). When this activation converges with activation sent from conceptual sources, a lexical item is selected and inserted in its proper slot. For the sentence "The bird flies", "bird" will be receiving activation from the message, but it will only become fully activated and selected for output when the proper syntactic-sequential state is reached.

Selection and insertion of lexical items by spreading activation from syntactic-sequential input has been emphasized in the Node Structure Theory of MacKay (1982, 1987). According to this theory, the syntactic-sequential states are represented as activated localist units that bear a close relation to the nodes of a syntactic tree. Other production theories also hold that these states represent direct implementations of syntactic structures (Berg, 1988; Dell, Burger, et al., 1997; Ferreira, 1996; Schade, 1992; Stemberger, 1985; Vosse & Kempen, 2000), but this is not the only way that they can be characterized. Syntactic-sequential states could, instead, correspond to the hidden-unit activation patterns of a simple recurrent network that learns to be sensitive to the distribution of words in utterances (Elman, 1990, 1993; Elman et al., 1996; Tabor & Tanenhaus, 1999). (See also Chang, 2002; Chang, Dell, Bock, & Griffin, 2000; Dell, Chang, & Griffin, 1999 for examples of this approach in production models.)

Thus, we see that the spreading activation models of lexical access in production allow for lexical retrieval to be governed by convergent activation from conceptual-semantic units and from syntactic-sequential states. Our claim is that the relative influence of these two activation sources can be explained by principles of learning, and that the resulting differences in syntactic and semantic influence help to explain key dissociations in aphasic lexical retrieval deficits.

1.2. Dissociations in aphasia

Since the earliest descriptions of aphasia, two broadly contrastive patterns of expression have been observed. So-called "non-fluent" aphasia is characterized by hesitant and effortful speech production with telegraphic-sounding sentences or sentence fragments, whereas "fluent" aphasics produce flowing grammatical speech replete with word substitutions which distort or decrease the meaning of the utterance. Jakobson (1956) described this dichotomy as a contrast between a syntagmatic disturbance, an impairment in the sequential ordering of words, and a paradigmatic disturbance, an impairment in the selection of appropriate words, a characterization which essentially captures the distinction between syntactic and semantic processing. Such broadly defined clinical categories have been criticized for their vagueness and heterogeneity (e.g., Schwartz, 1984) and the subjective nature of their definition (e.g., Gordon, 1998). However, the contrasting of maximally different patterns of performance within specific domains remains a valid method of exploring how language is represented in the brain.

One pattern of aphasic language production which has been the subject of considerable controversy concerning its underlying deficit is agrammatism. Agrammatic speech production is one of the factors contributing to non-fluency in aphasia, and is characterized by a tendency to omit closed-class (function) words and grammatical morphemes, and a reduction in the syntactic complexity and length of utterances (Berndt, 2001; Goodglass & Kaplan, 1983; Kolk & Heeschen, 1992; Rochon, Saffran, Berndt, & Schwartz, 2000; Saffran, Berndt, & Schwartz, 1989). Because the present study focuses on lexical access, the lexical retrieval deficits characteristic of agrammatism are contrasted with a different pattern of lexical retrieval deficits that has been observed in anomic aphasia. Whereas agrammatic aphasics show particular difficulty producing grammatical function words relative to content words, anomic aphasics demonstrate the opposite pattern (Goodglass & Kaplan, 1983).

That this dissociation occurs in both directions (i.e., function words better than content words in some subjects, content words better than function words in others) is of particular interest to investigators of language production, because it cannot be explained solely on the basis of the relative difficulty of the two types of items. Rather, such *double dissociations* have been interpreted as strong evidence for the autonomy of representational stores or processing modules. In the case of content and function words, for example, the double dissociation observed in aphasic performance supports the hypothesis discussed earlier, that the two types of words are accessed through different processing routes. Because agrammatism disrupts the structural frame, access to function words is more affected than access to content words (e.g., Bradley, Garrett, & Zurif, 1980; Segalowitz & Lane, 2000). Function words are also much more frequent than content words, which could explain their relative ease of production in normal speech and anomic aphasic speech, but leaves the opposite dissociation in agrammatism unexplained (Dell, 1990).

More specific double dissociations have also been noted within the category of content words. Several studies have illustrated that, whereas some aphasic patients (particularly anomics) have more difficulty producing nouns than verbs, others (particularly agrammatics) have relatively more difficulty producing verbs (Bates, Chen, Tzeng, Li, & Opie, 1991; Berndt, Haendiges, Mitchum, & Sandson, 1997; Berndt, Mitchum, Haendiges, & Sandson, 1997; Bird & Franklin, 1995/1996; Breen & Warrington, 1994; Caramazza & Hillis, 1991; Damasio & Tranel, 1993; Danieli, Giustolisi, Silveri, Colosimo, & Gainotti, 1994; Hillis & Caramazza, 1995; Luzzatti et al., 2002; Miceli, Silveri, Villa, & Caramazza, 1984; Williams & Canter, 1982, 1987; Zingeser & Berndt, 1988; Zingeser & Berndt, 1990; see Druks, 2002 for a review). Because nouns and verbs are both content words, this dissociation is not explained by the different retrieval processes described above, but several other hypotheses have been put forward. It has been suggested that verbs are generally more difficult to access, and that this difference is especially evident in naming tasks, in which the imageability of the word plays a significant role (Bird, Howard, & Franklin, 2000; Kohn, Lorch, & Pearson, 1989; Williams & Canter, 1987). This hypothesis is supported by findings of specific verb impairments in both fluent and non-fluent aphasics, while specific noun impairments have been reported less consistently, and primarily for fluent aphasics (e.g., Jonkers & Bastiaanse, 1998; Luzzatti et al., 2002; Zingeser & Berndt, 1990). Others have hypothesized that verb deficits are related to the greater grammatical complexity of verbs relative to nouns (e.g., Lapointe, 1985; Miceli et al., 1984; Saffran, Schwartz, & Marin, 1980; Zingeser & Berndt, 1990), a factor which affects agrammatic speakers more strongly than anomic speakers. However, neither of these explanations accounts for the occurrence of selective noun impairments.

The simplest and most popular explanation for the noun/verb double dissociation is that there is a transparent association of specific grammatical class deficits to disruptions in corresponding components of a grammatically subdivided lexicon (e.g., Caramazza & Hillis, 1991; Miceli et al., 1984; Williams & Canter, 1987). A related hypothesis postulates that the dissociation arises from a closely corresponding distinction between different types of semantic-conceptual representations. Specifically, disruption in the processing of perceptual features affects nouns more than verbs, whereas a deficit in functional features affects verbs more than nouns (Bates et al., 1991; Bird et al., 2000). This proposal carries the advantage of neurobiological plausibility, because verb (or functional-feature) impairments are usually associated with anterior brain lesions, which lie close to motor cortex, while noun (or perceptual-feature) impairments are more often associated with posterior lesions, which lie closer to sensory association areas (Bates et al., 1991; Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996; Peterson, Posner, Fox, Mintun, & Raichle, 1988). (A similar neurobiological explanation has been proposed to account for the content/function word dissociation, Pulvermüller, 1995.) Furthermore, simulated deficits in functional and perceptual features have been demonstrated to account, not only for the noun/verb dissociation, but also for a frequently observed semantic category dissociation between animate and inanimate objects (Bird et al., 2000, but see Caramazza & Shelton, 1998 for a critical review of this hypothesis).

Among patients with noted verb impairments, many of whom are agrammatic, recent studies have shown a further double dissociation between *light* and *heavy* verbs (Bencini & Roland, 1996; Berndt, Haendiges, et al., 1997; Berndt, Mitchum, et al., 1997; Breedin et al., 1998; Kohn et al., 1989). Light verbs are relatively unspecified semantically; they represent core predicates which often make up part of the semantic specification of a heavier verb (Jesperson, 1965; Kegl, 1995). For example, a light verb like "go" consists of semantic features (i.e., to move from one place to another) which form part of the definition of heavy verbs such as "fly" (i.e., to move from one place to another through the air) and "run" (i.e., to move from one place to another rapidly, by springing steps). Berndt, Mitchum, et al. (1997) noted that verb-impaired aphasic subjects tended to rely on light verbs in sentence production and story retelling tasks. The authors hypothesized that this might constitute a compensatory strategy in the face of failure to retrieve a specific verb (see also Kegl, 1995). Kohn and colleagues (1989) found similar tendencies in the sentences generated by both aphasic and normal control subjects. By contrast, Bencini and Roland (1996) compared agrammatic aphasic speech samples (Menn & Obler, 1990) to speech samples from unimpaired subjects and noted that the agrammatics used fewer light verbs than the controls. A similar dissociation has been shown between a greater-than-normal percentage of high frequency verbs in the spontaneous speech of a fluent aphasic subject and a preponderance of low-frequency verbs in an agrammatic aphasic subject (Bird & Franklin, 1995/1996).

Breedin and colleagues (1998) explicitly investigated the use of heavy and light verbs in a group of eight non-fluent aphasic subjects. First, subjects were tested to determine the degree to which agrammatism was evident in their spontaneous discourse (a story telling task), as indicated by four measures—the proportion of words in sentences, the ratio of nouns to verbs, the proportion of closed-class words, and the complexity of verb inflection used. On this basis, three subjects were classified as agrammatic (see Saffran et al., 1989 for details about the diagnostic procedure). In addition to these characteristics, all of the agrammatic subjects produced more heavy than light verbs, whereas all but one of the non-agrammatic subjects produced more light than heavy verbs.

The most striking contrast, and the most relevant to the present study, is illustrated by subject SS, who displayed all four indices of agrammatism, and subject VP, who displayed none, and has been classified as anomic (Breedin & Martin, 1996; Dell, Schwartz, et al., 1997). According to the number of words per minute produced in narrative, both subjects were less fluent than normal (mean 132 wpm), although subject SS (17 wpm) was far less fluent than subject VP (69 wpm). In addition, SS produced more than six times as many nouns as verbs (noun:verb ratio = 6.4), whereas VP, like the control subjects, produced almost equivalent numbers of each (noun:verb ratio = 1.2). Of the verbs produced in the narrative task, 85% of VP's were light verbs, but only 9% of SS's verbs were light. In a sentence completion task targeting particular verbs, VP also produced more simple, light, and general verbs than did SS, as a proportion of the verbs produced correctly by each subject. Among the errors produced on this task, SS tended to substitute heavy verbs for light verbs (17 heavy-for-light substitutions; 5 light-for-heavy substitutions), while VP tended to substitute light verbs for heavy verbs (5 heavy-for-light; 29 light-for-heavy).

These results confirm previously noted differences between agrammatic and anomic aphasic patients in both the degree of fluency exhibited (Goodglass & Kaplan, 1983) and in relative proportions of nouns and verbs produced (Williams & Canter, 1987). Furthermore, they extend the behavioral dissociation to the relative proportions of heavy and light verbs produced by agrammatics and anomics. To explain the heavy-verb advantage in agrammatic production, Breedin and colleagues (1998) hypothesized that, in some patients, the semantic richness (i.e., greater number of semantic features) of heavy verbs may make them more resistant to disruption than light verbs. Another possibility presented was that heavy verbs, being more constrained in the contexts in which they can occur, have representations which are less variable than those of light verbs, rendering them easier to retrieve consistently. To explain the opposite dissociation observed in some of the non-agrammatic subjects, Breedin and colleagues speculate that patients who show a light verb advantage may simply be taking advantage of a strategic reliance on higher frequency verbs.

Within a modular view of lexical representation, such a dissociation suggests separate representations for light and heavy verbs. However, whereas the distinctions between content and function words, and between nouns and verbs, are independently motivated on linguistic grounds, it is not clear why heavy and light verbs should dissociate. They belong to the same syntactic category and occupy similar semantic spaces, but differ quantitatively along several dimensions: light verbs, by definition, are specified by fewer semantic features than are heavy verbs; they are also less constrained by semantic context and, consequently, occur more frequently than heavy verbs. Furthermore, it offends the rule of parsimony to continue to fractionate the lexicon into increasingly specific modules in order to account for individual variations in aphasic performance. What is required instead is a solution that can account for the dissociation based on established featural and distributional differences between light and heavy verbs.

1.3. Dissociations and division of labor in connectionist models

One of the most compelling accomplishments of connectionist models that use learning to set connection weights is their ability to characterize the processing associated with quasi-regular linguistic domains. In such domains, the mappings that must be achieved to transform one type of representation into another can vary from perfectly systematic to completely unsystematic. Multi-component connectionist models can learn such regular, quasi-regular, and irregular relationships by dividing up the labor of such mappings among the different components. Consider connectionist models that learn to transform orthographic inputs to phonological outputs (as in reading aloud). These models allow for outputs to be achieved via two pathways—a direct pathway from phonological input, and an indirect pathway mediated by semantics (Harm, 1998; Harm & Seidenberg, 1999; Harm & Seidenberg, submitted; Plaut et al., 1996; Seidenberg & McClelland, 1989). Each word's computed phonology is influenced by both pathways, but to a greater or lesser extent, depending on the regularity of the word's spelling. Similarly, in the mapping from present to past form in English, there are forms (e.g., "take" → "took") that do not follow the typical pattern of analogously spelled words (e.g., "baked" → "baked", "rake" → "raked"). Joanisse and Seidenberg (1999) have shown that past tense forms of non-word verbs rely on input from a phonological component, whereas irregular past tense forms are more dependent on input from a semantic component.

The varying contributions of semantic, phonological, and orthographic knowledge to regular and irregular items result from the model's exposure to properties of the relevant mappings, together with an error-driven learning rule that creates cue competition. Consider spelling-to-sound mappings. For most words, the mapping from orthography to phonology is systematic, and therefore relatively easy for the model to learn. However, the mapping from semantics to phonology is almost always arbitrary, and this pathway is, therefore, at a comparative disadvantage. Because of the error-driven learning algorithm used in these models, the pathways compete. To the extent that the orthographic-phonological pathway is very successful in predicting the correct pronunciation, there is little learning in the semantic pathway. Thus, the orthographic-phonological pathway becomes specialized for the regular items (e.g., "soon"). In addition, this pathway is essential for non-words (e.g., "moop"), which have no semantic representations and must be pronounced by analogy to similarly spelled words. For irregular items (e.g., "flood"), on the other hand, the orthographic-phonological pathway is much less successful, so they must rely heavily on the semantically mediated pathway in order to compute pronunciations accurately.

The competition between cues or pathways is a natural consequence of error-driven learning. Error-driven algorithms such as the *delta rule* (Widrow & Hoff, 1988) only affect connection weights to the extent that output activations differ from target activations. A single input that is predictive of an output will acquire a strong connection to that output, but two inputs that are equally predictive of it will each acquire smaller weights because they do the job together. Such cue competition effects are common in learning. Overshadowing effects in Pavlovian conditioning (Kamin, 1969), illustrate that strong contingencies between a conditioned and an unconditioned stimulus may not be learned if other stimuli already predict the unconditioned stimulus. The principal model developed to account for these effects (Rescorla & Wagner, 1972) is actually formally identical to the delta rule (Sutton & Barto, 1981).

The division of labor among competing components in connectionist models has consequences for the application of those models to neuropsychological dissociations in language production. Traditionally, double dissociations are assumed to be transparently related to dissociations in the underlying functional architecture of the language processing system. With connectionist models, however, it is possible to obtain non-transparent double dissociations (Plaut, 1995). For example, the double dissociation between regular and irregular past tense formation that has been observed in brain-damaged patients (e.g., Ullman et al., 1997) is, in a traditional modular framework, hypothesized to arise from separate impairments to distinct processes—a process of rule application for regular past tense formation, and a lexical look-up process for irregular past tense formation (e.g., Pinker & Prince, 1988). However, Joanisse and Seidenberg's (1999) connectionist simulation of past tense formation exhibited this double dissociation without separate mechanisms for regular and irregular items. Because the model's phonological and semantic representations divided the labor of retrieving past tense forms as described above, with regular (non-word) forms more dependent on phonological weights and irregular forms on semantic weights, lesions to the phonological and semantic components in the model produced different patterns of results. Similar non-transparent double dissociations have been reproduced in connectionist models of reading, with respect to deficits in reading concrete and abstract words (Plaut & Shallice, 1993) and in reading regular and irregularly spelled words (Plaut et al., 1996).

The current study uses a similar approach to the observed dissociation between light and heavy verb production in aphasia. We start with the uncontroversial premise that lexical access depends on both conceptual-semantic and syntactic-sequential input. We then show that the relative contributions of each type of input can be determined by connectionist learning principles. This results in a continuum of dependence of lexical production on syntax and semantics, with function words at the syntactic end, and nouns and heavy verbs toward the semantic end. Light verbs fall at an intermediate point along the continuum. It is this trade-off between syntactic and semantic cues that results in the light/heavy verb dissociation, rather than their representation in distinct modules.

In the following sections, we describe four simulation studies. The first presents a model of normal lexical access in sentence production that learns to map two sources of information—a semantic-conceptual message and a sequence of syntactic states—onto an ordered sequence of words. In the second study, this model is lesioned in an attempt to account for aphasic lexical retrieval deficits, notably the double dissociation between light and heavy verbs. The third study manipulates the dimensions of difference between light and heavy verbs, in order to examine why the model produces this dissociation. In the final study, the performance of the model is tested in a new situation—producing single words instead of sentences—to determine whether the model can account for differences in aphasic performance on sentence production and naming tasks. In each study, the connection weights are examined to determine why the model behaves as it does; this direct interpretation is allowed by the simplicity of the model's architecture and vocabulary.

2. Study 1: normal sentence production

2.1. Network architecture

A model of sentence production was constructed from a two-layered non-linear feed-forward network. Fig. 1 shows all of the input and output units and the intended mappings between them. The output layer consisted of lexical nodes for each of the words in the sentences, and the input layer consisted of nodes for syntactic-sequential states and semantic features defining each lexical item. There were no hidden layers. To keep the network manageable and interpretable, only intransitive sentences were modeled, and the only syntactic nodes included represented the three syntactic states corresponding to the initial determiner, the subject noun, and the verb. The only semantic features included were those sufficient to distinguish each word from the others in the network. Lexical items were uninflected, also for the sake of simplicity, because inflectional distinctions were not important for the purposes of this study. Furthermore, the model is not intended to provide a data-fitting mechanism, but to better understand how dissociations in behavior may arise in a system. The small vocabulary and simple structure

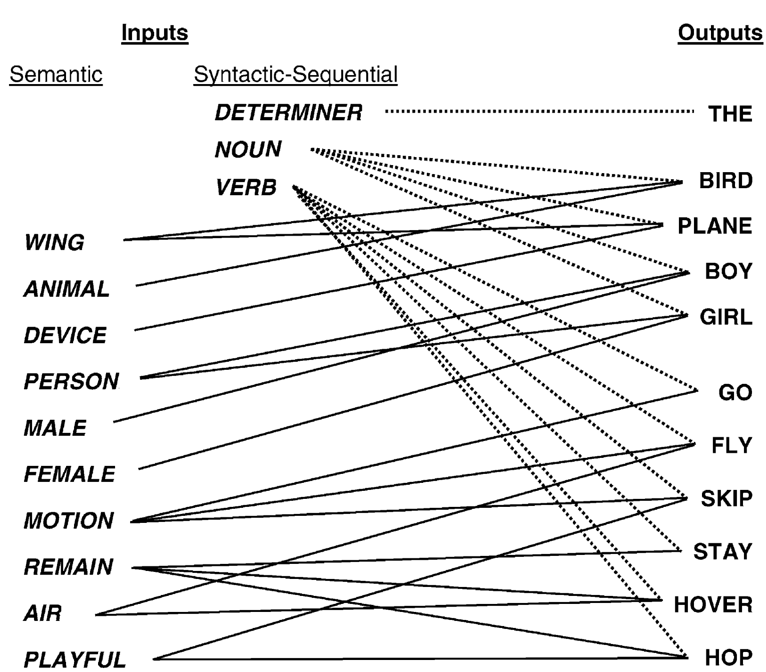

Fig. 1. The architecture of the network. Only intended mappings are shown, although potential connections exist between all input and output nodes. Syntactic connections are shown by dotted lines; semantic connections by solid lines.

of the model allowed a straightforward interpretation of its output, in relation to its input and various properties.

The network was set up to reflect the relative properties of light and heavy verbs in natural language samples. The two light verbs were specified by fewer semantic features than the heavy verbs, and these features constituted a subset of the features specifying the heavy verbs. Thus, GO was specified by the feature *MOTION*, FLY was specified by the features *MOTION* and *AIR*, and SKIP was specified by the features *MOTION* and *PLAYFUL*. Light verbs were represented in target sentences more frequently than heavy verbs, and had fewer contextual constraints. That is, they co-occurred with more of the model's nouns—or were more widely distributed—than the heavy verbs: BIRDs, PLANEs, BOYs and GIRLs can GO and STAY, but only BIRDs and PLANEs can FLY or HOVER, and only BOYs and GIRLs can SKIP or HOP.

2.2. Training the model

All input nodes were connected to all output nodes, but the relative strengths of these connections were set by training the model to map semantic messages onto grammatical sequences of words. The input consisted of sets of features defining sixteen simple intransitive sentences, and the desired output consisted of the lexical items appropriate to each sentence. For example, the target sentence "The bird flies" was trained by the following three input–output sets:

Input		Output
Syntactic-sequential state	Semantic features	
DETERMINER	*WING, ANIMAL, MOTION, AIR*	THE
NOUN	*WING, ANIMAL, MOTION, AIR*	BIRD
VERB	*WING, ANIMAL, MOTION, AIR*	FLY

Note that the semantic message remains static, but the syntax changes throughout the sentence to ensure that the correct lexical item is output at the correct time. By keeping the semantics static, the model reflects a key assumption of modern spreading activation theories of production (e.g., Berg, 1988; Chang et al., 2000; Dell, Burger, et al., 1997; MacKay, 1987; Schade, 1992; Stemberger, 1985). That is, the message is largely in place when production begins (see Griffin & Bock, 2000), and the part of the model that changes to create ordered output is the syntactic-sequential state. This aspect of the model will be important in explaining the results of Study 4, the comparison of single-word naming with sentence production.

The model was constructed and trained using "tlearn" software (Plunkett & Elman, 1997). The learning algorithm used was the delta rule [$\Delta w_{io} = \varepsilon f'$ (net) $(t_o - a_o)a_i$] (Widrow & Hoff, 1988). The change in connection weight (Δw_{io}) between a given output node (a_o) and a given input node (a_i) was determined by the difference between the desired, or target, activation of the output node (t_o) and its actual activation (a_o), multiplied by its input activation (a_i). Output activations were a logistic function of their net inputs (ranging from 0 to 1); f' (net) refers to the slope of this activation function at the net input ($a_o(1 - a_o)$). The learning rate used was small ($\varepsilon = 0.1$), in order to allow weights to be changed gradually as they approach a solution with each successive trial. Initial weights were randomly assigned to values between -0.1 and $+0.1$. There were no biases on output units. The model was considered to be "trained" once the root mean squared (RMS) error, representing the collective discrepancy between desired and actual activations of targeted nodes, dropped below a criterion value of 0.50.

2.3. The model's output

After 5,000 sweeps through the training corpus of 16 sentences, the model reached a RMS error of 0.407. At this point, the model was run once on each of the 16 sentences to assess whether the input features resulted in the desired output sentences. For each sentence, the model produced three sets of activation levels, one for each sequential state within the sentence. At each sequential state, the accuracy of the intended target word was calculated by a stochastic choice rule (McClelland & Rumelhart, 1981), which determined the probability of production of a lexical item given its activation level relative to the activation levels of all the lexical items. (In this determination, activations (a) were transformed to $e^{\mu a}$, with $\mu = 10$.) These probabilities were used to determine the model's actual output for each target sentence and, when targets were not correctly produced, the errors most likely to be produced instead.

Table 1 shows the activation levels and production probabilities for each lexical item in two sample sentences, one with a light verb and one with a heavy verb. To illustrate, in the sentence "The plane flies" in the unlesioned model, the probability that "plane" would be produced is calculated by comparing its activation level ($a = 0.741$) to the activations of all the other possible outputs, yielding a probability of .988. Because the model contains no hidden units, the error-driven learning mechanism (the delta rule) should converge on the best set of weights. That is, there is a single global error minimum, a fact which was confirmed by two replications of the model. Because the three runs of the model showed the same results, only one run, randomly sampled, is reported here.

Production probabilities for each lexical type (determiner, nouns, heavy verbs, and light verbs) were then averaged across all the target sentences. These accuracy levels reflect how predictable a particular type of lexical item is, based on its input semantic and syntactic cues and on how well the model has learned these cues during the training phase. The determiner THE, which appeared at the beginning of every sentence, was extremely well learned, with a mean probability of production across sentences of .998. The four nouns, evenly divided among the training sentences, were also very well learned, with a mean production probability of .989. Light verbs occurred with the same frequency as the nouns and showed a mean probability of production of .954. This accuracy level is slightly below that for nouns, because the semantic cues for light verbs are less predictable; they have no semantic features that do not also apply to heavy verbs. Relative to heavy verbs, however, which showed a mean production probability of .809, light verbs were learned more quickly and considerably more reliably. This is because they occurred twice as often as the heavy verbs, and were therefore given more practice during the training phase. A one-way ANOVA, treating sentences as the random effect, showed that the part of speech (determiner, noun, light verb, or heavy verb) had a significant effect on probability of production ($F_{(3,44)} = 67.6$, $p < .001$). Scheffé *post hoc* pairwise comparisons ($p < .001$) showed that heavy verbs were less likely to be produced correctly than all other parts of speech, and that determiners had a higher probability of production than light verbs.

Table 1
Output activations and resulting production probabilities for two sample sentences: unlesioned and lesioned models

THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP	Target	*p*
Unlesioned model												
The bird goes												
0.934	0.120	0.008	0.005	0.006	0.125	0.056	0.027	0.009	0.015	0.008	the	.999
0.055	0.737	0.143	0.102	0.108	0.116	0.053	0.027	0.009	0.015	0.008	bird	.987
0.056	0.121	0.008	0.005	0.006	**0.672**	0.259	0.120	0.115	0.075	0.035	go	.963
The plane flies												
0.925	0.007	0.127	0.003	0.003	0.024	0.180	0.015	0.002	0.040	0.005	the	.998
0.049	0.131	**0.741**	0.059	0.057	0.022	0.172	0.015	0.002	0.040	0.005	plane	.988
0.049	0.007	0.119	0.003	0.003	0.257	**0.564**	0.072	0.025	0.177	0.020	fly	.899
Syntactically lesioned model												
The bird goes												
0.311	0.427	0.044	0.029	0.029	0.366	0.185	0.088	0.037	0.050	0.026	the	.149
0.311	**0.427**	0.044	0.029	0.029	0.366	0.185	0.088	0.037	0.050	0.026	bird	.476
0.311	0.427	0.044	0.029	0.029	**0.366**	0.185	0.088	0.037	0.050	0.026	go	.259
The plane flies												
0.283	0.039	0.440	0.016	0.015	0.089	0.456	0.052	0.008	0.123	0.015	the	.082
0.283	0.039	**0.440**	0.016	0.015	0.089	0.456	0.052	0.008	0.123	0.015	plane	.392
0.283	0.039	0.440	0.016	0.015	0.089	**0.456**	0.052	0.008	0.123	0.015	fly	.460
Semantically lesioned model												
The bird goes												
0.969	0.155	0.157	0.154	0.165	0.199	0.208	0.223	0.189	0.228	0.237	the	.996
0.115	**0.790**	0.785	0.791	0.803	0.185	0.199	0.221	0.189	0.228	0.238	bird	.243
0.116	0.156	0.147	0.147	0.166	**0.780**	0.607	0.586	0.773	0.607	0.573	go	.390
The plane flies												
0.969	0.155	0.157	0.154	0.165	0.199	0.208	0.223	0.189	0.228	0.237	the	.996
0.115	0.790	**0.785**	0.791	0.803	0.185	0.199	0.221	0.189	0.228	0.238	plane	.231
0.116	0.156	0.147	0.147	0.166	0.780	**0.607**	0.586	0.773	0.607	0.573	fly	.069

Note. Highlighted numbers indicate activation levels of the targeted output nodes.

Production probabilities of non-target items were also examined to determine the main competitors for each target. For the noun slot in each sentence, the next most likely item to be produced was a semantically related noun (e.g., PLANE for BIRD), and in the verb slot, the next most likely item was a semantically related verb (e.g., GO for FLY). (This can be seen by looking at their relative activation levels in Table 1.) Thus, in the unlikely event of an error, syntactic category constraints, which have been observed in studies of both normal and aphasic errors (e.g., Gagnon, Schwartz, Martin, Dell, & Saffran, 1997; Garrett, 1992), were preserved. Given the observed probabilities, the most likely error to occur would be the substitution of a light verb for a heavy verb, replicating findings from the spontaneous speech of normal speakers (e.g., Kohn et al., 1989).

The differential performance of light and heavy verbs can be further clarified by examining the relative weights on syntactic and semantic connections. Table 2 shows a matrix of the trained connection weights between all input and output nodes. When targeted for production, heavy verbs receive a small positive input from the syntactic *VERB* node (e.g., the connection weight to FLY is 0.43), but most of their activation comes from the semantic feature nodes, particularly their most specifying semantic feature (the connection from *AIR* to FLY has a large positive weight of 1.22). At the same time, specifying semantic features like *AIR* send strong negative inputs to inhibit the activation of other verbs (the connection from *AIR* to GO carries a weight of -1.73). In contrast, light verbs are primarily activated by the syntactic *VERB* node (the weight to GO is 1.27), and to a much lesser extent by their semantic features (0.50 from *MOTION* to GO). Thus, the expected division of labor is observed between syntactic and semantic cues. During the training phase, this division of labor comes about because there are no unique semantic predictors for the light verbs. Therefore, the syntax must assume a greater proportion of the responsibility for their production. This pattern of inputs, along with their greater frequency of occurrence, serves to make the light verbs function to some extent as default verbs. On the other hand, the semantic feature *AIR*, in combination with the *MOTION* feature, is a reliable predictor for the production of FLY. A strong connection is established between *AIR* and FLY, in order to overcome the tendency for the *VERB* node, which also becomes activated when FLY is the target, to activate GO instead.

The other types of lexical items also show the predicted division of labor between syntax and semantics. The determiner has no semantic features in this model, so necessarily relies on strong positive input from its syntactic-sequential specification. Like the heavy verbs, the nouns rely largely on their unique semantic features. (Like light verbs, nouns also receive a significant proportion of activation from their syntactic node; however, the greater importance of the semantic features to noun production will become evident when the model is lesioned.) While the distributional and frequency characteristics of the lexical items determine the optimum division of labor for their accurate and efficient production under normal conditions, this trading off of cues also makes them more or less susceptible under different conditions of damage.

3. Study 2: lesioning the model

Once the “normal” model was sufficiently trained and examined, it was lesioned to simulate either a syntactic (agrammatic-like) deficit or a semantic (anomic-like) deficit. Syntactic lesions

Table 2
Input–output connection weights: original model

Input nodes	Output nodes									
	THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP
DETERMINER	**3.44**	−1.70	−1.68	−1.70	−1.62	−1.39	−1.34	−1.25	−1.46	−1.22	−1.17
NOUN	−2.04	**1.33**	**1.29**	**1.33**	**1.41**	−1.48	−1.39	−1.26	−1.46	−1.22	−1.16
VERB	−2.03	−1.69	−1.76	−1.76	−1.61	**1.27**	**0.43**	**0.35**	**1.23**	**0.43**	**0.29**
MOTION	−0.40	−1.05	−1.04	−1.04	−1.02	**0.50**	**−0.37**	**−0.35**	−2.05	−1.87	−1.67
REMAIN	−0.35	−1.04	−0.99	−1.02	−1.04	−1.96	−1.94	−1.70	**0.47**	**−0.35**	**−0.35**
AIR	−0.11	−0.03	−0.10	−0.54	−0.59	−1.73	**1.22**	−0.49	−1.64	**1.02**	−0.49
PLAYFUL	0.00	−0.61	−0.62	−0.08	−0.04	−1.69	−0.50	**1.12**	−1.68	−0.48	**0.91**
WING	−0.25	**−0.45**	**−0.45**	−1.66	−1.69	−0.77	−0.74	−1.29	−0.80	−0.75	−1.40
ANIMAL	−0.14	**1.21**	−1.59	−0.80	−0.80	−0.28	−0.37	−0.70	−0.42	−0.32	−0.56
DEVICE	−0.16	−1.68	**1.35**	−0.86	−0.90	−0.33	−0.29	−0.78	−0.38	−0.36	−0.62
PERSON	−0.31	−1.58	−1.67	**−0.42**	**−0.50**	−0.70	−1.41	−0.71	−0.71	−1.39	−0.72
MALE	−0.17	−0.89	−0.81	**1.28**	−1.68	−0.38	−0.73	−0.38	−0.34	−0.70	−0.33
FEMALE	−0.14	−0.88	−0.82	−1.66	**1.28**	−0.35	−0.67	−0.34	−0.35	−0.64	−0.45

Note. Highlighted numbers indicate activated input features for each target output.

involved setting all the connection weights between syntactic input features and output lexical items to zero, effectively severing all syntactic connections. Semantic lesions were simulated by severing all connections between semantic input features and output lexical items. Partial lesions were also simulated, by systematically reducing the weights from either syntactic or semantic nodes. (Note that the same effects could have been achieved by eliminating or reducing the activation levels of the input nodes.)

In the syntactically lesioned model, all word-order cues were lost, so the output activation pattern across lexical items did not change throughout the three sequential states in a given sentence (see Table 1 for examples). Averaged across sentences, nouns were most accurately produced (mean probability $= .499$), followed by heavy verbs (mean probability $= .354$), then light verbs (mean probability $= .208$). The determiner had consistently low activation levels, with a mean production probability of about .123. Scheffé *post hoc* tests showed that each pairwise comparison was significant: nouns had a higher probability of production than all other words ($p < .001$); determiners had a lower probability of production than both heavy verbs ($p < .001$) and light verbs ($p < .05$); and heavy verbs were more likely than light verbs to be produced correctly ($p < .001$).

These probabilities predict particular sentence patterns. Because of the extremely low probability of production of the determiner, both the noun and the verb (especially a heavy verb) were more likely than the determiner to be produced at the beginning of the sentence. Similarly, when a light verb was targeted, the sentence's noun was more highly activated. Thus, a typical sentence output might be "Bird bird flies", or "Bird bird bird". If we allow that low activation levels might sometimes result in no output at all, a property which could easily be built into the model using a minimum threshold of activation, then some words might be omitted altogether, resulting in one- or two-word sentences such as "Bird" or "Bird flies". (The relevance of such a threshold will be explored further in the discussion.) All of these production patterns are characteristic of agrammatism. Note that error substitutions come from within the target sentence (i.e., syntagmatic errors), such that the resulting output reflects clinically observed patterns of function word omission and sentence fragment production.

To simulate more realistic lesions, syntactic connection weights were reduced to 75%, 50% and 25% of their original values. In Fig. 2, the resulting mean probabilities of production for each lexical type are plotted at these three connection weight strengths, the completely lesioned connection weight strength (0%), and the unlesioned connection weight strength (100%). With syntactic lesions (Graph A), the light verb advantage of the unlesioned model is evident at the 100% connection strength point. As the severity of the lesion increases, however, this advantage is gradually reversed until a heavy verb advantage is evident. The actual point at which this reversal occurs is an artifact of the somewhat arbitrary quantitative characteristics of the model (e.g., numbers of semantic features, frequency of occurrence of lexical items, number of competitors in the lexicon). Nevertheless, the apparent relationship between the severity of the syntactic deficit and the direction of the dissociation may help to explain some of the variability among agrammatic patients. For example, as discussed earlier, Breedin and colleagues (1998) found a heavy/light verb dissociation in the spontaneous speech of only four of the five verb-impaired subjects in their study; the one who did not show the dissociation was the one with the least severe verb impairment.

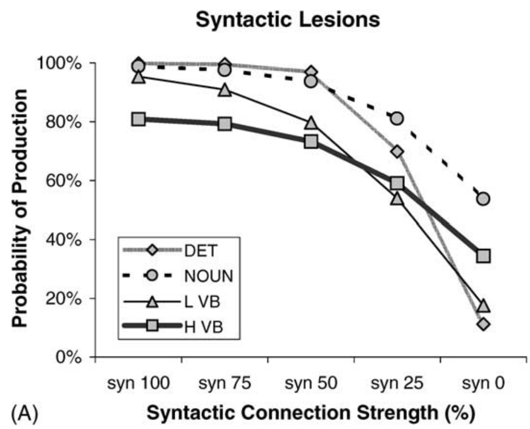

Fig. 2. Effect of lesions on probability of production: sentence production. To simulate partial lesions, connections between output nodes and either syntactic-sequential inputs (A) or semantic inputs (B) were systematically reduced to 75%, 50%, and 25% of their original strength. Unlesioned (100%) and completely lesioned (0%) models are also included for reference.

In the semantically lesioned model, the output activation pattern varied across sequential states, but because all semantic information was lost, the same three activation patterns occurred for each target sentence (see Table 1). The determiner, which in this simulation has no semantic features, retained a high likelihood of production (mean probability = .996). The four nouns were all highly activated, but to virtually the same degree, giving each about a .25 probability of production, no matter which one was the target. Light verbs had a much higher likelihood of production (mean probability = .377) than heavy verbs (mean probability = .061), such that light verbs were likely to substitute for heavy verb targets. All *post hoc* pairwise comparisons were significant ($p < .001$). The most probable output would therefore take the form of a semantically empty sentence such as "The [randomly chosen noun] goes" (or, if such a default light noun had been included in the network, "The *thing* goes"), a typically anomic response. Because the most likely errors are substitutions from within the same grammatical category (i.e., paradigmatic errors), syntactic structure remains accurate and complete, although the meaning of the utterance is disrupted.

Graph B in Fig. 3 shows the effect of partial semantic lesions. The "normal" dissociation (i.e., the light verb advantage) is exaggerated by semantic lesions of increasing severity, a finding which accords well with studies of anomia in normal aging and dementia (Bowles & Poon, 1985; Nicholas, Obler, Au, & Albert, 1996). It is also of note that nouns, having the same number of semantic features as heavy verbs, follow roughly the same slope with severity. Determiners, on the other hand, which have no semantic features, show virtually no change.

As predicted, the division of labor between syntactic and semantic information gives rise to characteristic anomic and agrammatic speech patterns, particularly in the production of light and heavy verbs. Reference back to Table 2 illustrates why. In syntactically lesioned models, weights from the *VERB* node are reduced or set to zero, which disrupts light verbs such as GO to a far greater extent than heavy verbs. In semantically lesioned models, reducing the semantic inputs forces targets to rely on their syntactic input. The difference in weights from the *VERB* node to GO and FLY shows how heavy verbs are disrupted more than light verbs when this happens. The factors promoting this division of labor were explored more deeply in the next study.

4. Study 3: manipulating the characteristics of light and heavy verbs

Several different simulations were run, modifying the original set of training sentences to systematically manipulate the variables that differentiate light from heavy verbs—frequency of occurrence, range of distribution, and number of semantic features. In real vocabularies, these variables are confounded: words with fewer semantic features are less constrained in terms of the contexts in which they can occur; being more widely distributed means that these items also occur more frequently. Different hypotheses have been put forth in the literature concerning which aspect of the difference between light and heavy verbs determines their differential behavior (Berndt, Haendiges, et al., 1997; Berndt, Mitchum, et al., 1997; Breedin et al., 1998). A model such as this affords the opportunity to manipulate each variable in turn, in order to partial out their effects.

For each version of the model, once the unlesioned model was trained, it was syntactically and semantically lesioned to assess how the double dissociation was affected. The following manipulations are described in reference to how they differ from the original model, which represents our implementation of the natural language situation. Table 3 shows the mean probabilities of production for each version of the model, in unlesioned and lesioned forms (for now, disregard the probabilities listed under "Naming Simulation"). The size of the dissociation in each models is indexed by the difference in probabilities between light and heavy verbs (e.g., .145 for the unlesioned original model). This difference is calculated according to the predicted direction, as shown in the original model. That is, heavy verbs are subtracted from light verbs where probabilities for light verbs are expected to be greater than for heavy verbs, and *vice versa*. So a negative dissociation indicates that the difference is in the direction opposite to that shown in the original model. The size of the *double* dissociation in the lesioned models is indexed by the combined difference scores of the lesioned models. So, for example, the dissociation in the syntactically lesioned original model is the probability of heavy verb production minus the probability of light verb production (.354 − .208 = .146), and the dissociation in the semantically lesioned model is .316, the light verb probability (.377) minus the heavy verb probability (.061). These combine to a double dissociation of .462. Note that a dissociation that is in a direction opposite to that predicted will reduce the size of the double dissociation.

4.1. The role of frequency

The most obvious potential source of the light verbs' advantage in the original model is their relatively high frequency of occurrence—they each occurred in four sentences, whereas each heavy verb only occurred twice. To assess the role played by this factor, the training corpus was altered to match the frequencies of each heavy verb to each light verb, while preserving the discrepancy in their range of distribution. In the resulting output, light verbs were significantly

Table 3
Mean probabilities of production across different versions of the model

Model	Sentence production					Naming simulation			
	DETERMINER	NOUN	LIGHT VERB	HEAVY VERB	DISSOCIATION	NOUN	LIGHT VERB	HEAVY VERB	DISSOCIATION
Mean probabilities of production (%)									
Original model									
Unlesioned	0.998	0.989	0.954	0.809	0.145	0.992	0.930	0.852	0.078
Syntactic lesion	0.123	0.499	0.208	0.354	0.146	0.901	0.662	0.843	0.181
Semantic lesion	0.996	0.249	0.377	0.061	0.316	0.249	0.377	0.061	0.316
					0.462				**0.497**
Equal Frequency model									
Unlesioned	0.998	0.990	0.813	0.936	−0.123	0.992	0.684	0.942	−0.258
Syntactic lesion	0.118	0.458	0.140	0.414	0.274	0.906	0.453	0.897	0.444
Semantic lesion	0.996	0.249	0.246	0.126	0.120	0.249	0.246	0.126	0.120
					0.394				**0.564**
Equal Distribution model									
Unlesioned	0.998	0.988	0.985	0.876	0.109	0.991	0.752	0.845	−0.093
Syntactic lesion	0.102	0.451	0.325	0.422	0.097	0.843	0.225	0.845	0.620
Semantic lesion	0.996	0.249	0.386	0.056	0.330	0.249	0.387	0.056	0.331
					0.427				**0.951**
Equal Frequency/Equal Distribution model									
Unlesioned	0.998	0.990	0.951	0.943	0.008	0.991	0.433	0.911	−0.478
Syntactic lesion	0.109	0.451	0.250	0.435	0.185	0.851	0.184	0.847	0.663
Semantic lesion	0.996	0.249	0.260	0.119	0.141	0.256	0.260	0.119	0.141
					0.326				**0.804**
Equal Features model									
Unlesioned	0.999	0.990	0.969	0.926	0.043	0.990	0.981	0.868	0.113
Syntactic lesion	0.122	0.450	0.333	0.372	0.040	0.937	0.933	0.793	−0.140
Semantic lesion	0.996	0.249	0.281	0.109	0.173	0.249	0.281	0.109	0.172
					0.212				**0.032**

Note. Highlighted numbers represent the size of the double dissociation between light and heavy verb production (see text for details).

less likely ($p < .001$) to be produced correctly (mean probability $= .813$) than were heavy verbs (mean probability $= .936$). Not only did the advantage for light verbs disappear, but it was reversed, resulting in a negative dissociation of $-.123$.

Examination of the new connection weights revealed that the effect of equalizing frequency was to bring the activation from input nodes shared by light and heavy verbs closer together. (For example, compare the connections from the features *VERB* and *MOTION* to the verbs GO and FLY in the weight matrix in Table 4 to the same weights in the original matrix in Table 2.) This reduces the light verb advantage from syntactic cues, and increases the heavy verb advantage from semantic cues. The combined effect was to reduce the double dissociation from .462 in the original model to .394.

4.2. The role of distribution

The heavy verb advantage revealed in the model with equal frequencies is probably attributable to differences in distribution between light and heavy verbs: the greater the variety of contexts in which a word appears, the more difficult it is to learn, as suggested by Breedin and colleagues (1998). To test this, the model was re-run on a new set of training sentences in which the range of distribution for light verbs was restricted to match that of heavy verbs. This manipulation had the effect of increasing the accuracy of both types of verbs in the unlesioned model (see Table 3), although the light/heavy dissociation is maintained ($p < .001$).

The effect of distribution on the division of labor is manifest primarily in the connections between verbs and the semantic features of the nouns with which they co-occur. These connections are naturally inhibitory, in order to prevent activation of a verb when a noun is required. However, connections from noun features are *less* inhibitory when they co-occur with a given verb. For example, in the original model (see Table 2), inhibition from the feature *PERSON* is twice as strong to FLY (-1.41) than to GO (-0.70), because the nouns specified by the *PERSON* feature (BOY and GIRL) never occur with the verb FLY. When the distribution of light verbs is artificially restricted, noun features gain importance in their ability to predict the occurrence of light verbs (see the weight matrix in Table 5). As a consequence, the heavy verb advantage is reduced. The light/heavy dissociation becomes smaller in the syntactically lesioned model, but slightly larger in the semantically lesioned model. Note that the semantically lesioned model shows very little change compared to the original model, because the effect of distribution is primarily on those connections which are severed. Overall, the double dissociation in the lesioned models (.427) was slightly less than in the original model (.462), reflecting the reduced heavy verb advantage.

When both the frequency and the range of distribution were equalized for heavy and light verbs, the network connections show combined effects of the two previously described manipulations (see Table 6). Given their equivalent distribution and frequency, the light/heavy verb dissociation in the unlesioned model is negligible. Thus, frequency of occurrence and distribution have opposing effects, but the effect of frequency appears to be stronger, at least in this implementation, conferring an overall production advantage on light verbs despite their wider distribution. Even when frequency and distribution are equalized, however, the remaining difference between light and heavy verbs—that is, the number of semantic features—still prompts a division of labor between syntactic and semantic cues, which

Table 4
Input–output connection weights: Equal Frequency model

Input nodes	Output nodes									
	THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP
DETERMINER	**3.41**	−1.71	−1.64	−1.61	−1.67	−1.18	−1.42	−1.44	−1.15	−1.47	−1.44
NOUN	−2.08	**1.32**	**1.39**	**1.42**	**1.36**	−1.17	−1.42	−1.44	−1.14	−1.46	−1.45
VERB	−2.08	−1.72	−1.64	−1.62	−1.68	**0.96**	**0.67**	**0.66**	**0.97**	**0.60**	**0.67**
MOTION	−0.30	−1.00	−1.00	−1.05	−1.00	**0.30**	**−0.10**	**−0.19**	−1.69	−2.05	−2.10
REMAIN	−0.30	−0.99	−1.00	−1.05	−1.00	−1.60	−1.97	−2.08	**0.22**	**−0.17**	**−0.18**
AIR	−0.12	−0.09	−0.08	−0.80	−0.81	−1.55	**1.04**	−0.76	−1.52	**1.00**	−0.70
PLAYFUL	−0.09	−0.69	−0.76	−0.08	−0.08	−1.53	−0.75	**1.03**	−1.60	−0.70	**1.05**
WING	−0.32	**−0.43**	**−0.48**	−1.65	−1.55	−0.72	−0.77	−1.38	−0.71	−0.68	−1.46
ANIMAL	−0.11	**1.32**	−1.65	−0.75	−0.84	−0.33	−0.40	−0.80	−0.33	−0.37	−0.76
DEVICE	−0.10	−1.65	**1.31**	−0.74	−0.85	−0.34	−0.41	−0.78	−0.33	−0.39	−0.72
PERSON	−0.36	−1.59	−1.58	**−0.39**	**−0.46**	−0.64	−1.49	−0.76	−0.71	−1.47	−0.73
MALE	−0.10	−0.83	−0.86	**1.23**	−1.66	−0.42	−0.77	−0.34	−0.29	−0.70	−0.38
FEMALE	−0.09	−0.80	−0.84	−1.73	**1.31**	−0.40	−0.78	−0.37	−0.28	−0.68	−0.38

Note. Highlighted numbers indicate activated input features for each target output.

Table 5
Input–output connection weights: Equal Distribution model

Input nodes	Output nodes										
	THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP
DETERMINER	**3.44**	−1.66	−1.67	−1.62	−1.68	−1.55	−1.23	−1.34	−1.56	−1.36	−1.21
NOUN	−2.02	**1.36**	**1.32**	**1.32**	**1.32**	−1.55	−1.21	−1.38	−1.57	−1.34	−1.18
VERB	−2.06	−1.68	−1.73	−1.72	−1.66	**1.45**	**0.40**	**0.53**	**1.39**	**0.49**	**0.39**
MOTION	−0.28	−1.04	−0.97	−1.14	−1.04	**−0.21**	**−0.38**	**0.00**	−1.49	−2.06	−1.72
REMAIN	−0.30	−1.11	−1.18	−1.01	−0.96	−1.60	−1.71	−2.11	**−0.13**	**−0.02**	**−0.43**
AIR	−0.09	0.01	0.02	−0.58	−0.60	−2.11	**1.35**	−0.51	−0.85	**0.69**	−0.52
PLAYFUL	−0.14	−0.65	−0.65	−0.03	−0.04	−0.81	−0.55	**0.61**	−2.06	−0.52	**1.32**
WING	−0.36	**−0.46**	**−0.46**	−1.58	−1.69	−0.06	−0.99	−1.73	−1.58	−0.54	−1.23
ANIMAL	−0.19	**1.27**	−1.64	−0.82	−0.84	−0.10	−0.38	−0.92	−0.84	−0.29	−0.61
DEVICE	−0.19	−1.67	**1.26**	−0.81	−0.82	−0.04	−0.40	−0.93	−0.86	−0.24	−0.55
PERSON	−0.36	−1.64	−1.57	**−0.37**	**−0.44**	−1.54	−1.25	−0.43	−0.11	−1.74	−0.88
MALE	−0.18	−0.77	−0.75	**1.27**	−1.74	−0.82	−0.55	−0.27	−0.11	−0.83	−0.41
FEMALE	−0.11	−0.77	−0.74	−1.64	**1.13**	−0.82	−0.56	−0.25	−0.16	−0.79	−0.47

Note. Highlighted numbers indicate activated input features for each target output.

Table 6
Input–output connection weights: Equal Frequency/Equal Distribution model

Input nodes	Output nodes										
	THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP
DETERMINER	**3.41**	−1.70	−1.63	−1.68	−1.66	−1.30	−1.42	−1.54	−1.29	−1.50	−1.40
NOUN	−2.09	**1.33**	**1.40**	**1.36**	**1.36**	−1.32	−1.42	−1.53	−1.30	−1.49	−1.38
VERB	−2.09	−1.71	−1.64	−1.69	−1.67	**1.07**	**0.60**	**0.76**	**1.09**	**0.80**	**0.62**
MOTION	−0.26	−0.97	−1.02	−1.02	−1.02	**−0.20**	**−0.21**	**0.09**	−1.40	−2.29	−1.99
REMAIN	−0.29	−1.07	−1.11	−0.95	−0.96	−1.34	−1.95	−2.23	**−0.20**	**0.04**	**−0.30**
AIR	−0.06	−0.04	−0.05	−0.71	−0.75	−1.99	**1.27**	−0.70	−0.86	**0.44**	−0.81
PLAYFUL	−0.16	−0.81	−0.83	−0.06	−0.08	−0.85	−0.75	**0.45**	−1.94	−0.67	**1.26**
WING	−0.34	**−0.42**	**−0.45**	−1.60	−1.60	−0.20	−0.93	−1.76	−1.52	−0.41	−1.17
ANIMAL	−0.16	**1.27**	−1.67	−0.84	−0.81	−0.05	−0.43	−0.85	−0.64	−0.17	−0.65
DEVICE	−0.17	−1.70	**1.29**	−0.82	−0.82	−0.06	−0.44	−0.82	−0.63	−0.17	−0.60
PERSON	−0.28	−1.51	−1.56	**−0.39**	**−0.41**	−1.46	−1.29	−0.38	−0.19	−1.75	−0.87
MALE	−0.14	−0.79	−0.78	**1.21**	−1.71	−0.75	−0.58	−0.21	−0.07	−0.88	−0.47
FEMALE	−0.14	−0.80	−0.75	−1.76	**1.23**	−0.73	−0.56	−0.22	−0.07	−0.86	−0.47

Note. Highlighted numbers indicate activated input features for each target output.

gives rise to dissociations in both lesioned models. The observed double dissociation was .326.

4.3. The role of semantic richness

In the final manipulation, the original model with unequal frequencies and distributions for light and heavy verbs was modified by the addition of a "neutral" semantic input node to each light verb. Thus, each of the light verbs was specified by two semantic features (GO by *MOTION* and *NEUTRAL*; STAY by *REMAIN* and *NEUTRAL*) just like each of the heavy verbs. The resulting output activations illustrate that the advantage for light verbs over heavy verbs was maintained, though it was not as great as in the original model (see Table 3). Furthermore, when frequency and distribution were equalized as described above, the only remaining difference being number of features, light and heavy verbs were equally accurate. Having fewer semantic features, then, does not seem to significantly affect the accuracy of production in normal speakers, that is, as long as both semantic and syntactic sources of information are available. However, this factor *did* show an effect on the behavior of the lesioned models.

Increasing the number of semantic features defining light verbs served to transfer much of responsibility for their activation to the new *NEUTRAL* feature (see Table 7). The feature *NEUTRAL* does for GO and STAY what *AIR* does for FLY and HOVER. As a consequence, GO is no longer automatically activated as a subset of the features of FLY. The *VERB* node also becomes less important in activating GO, because the labor is divided among three input nodes rather than two. Because of their increased semantic richness, light verbs are not as devastated by a syntactic lesion, so the heavy verb advantage is almost completely eradicated. However, because light verbs retained their frequency advantage, they also retained a small but significant advantage over heavy verbs in the unlesioned and semantically lesioned models. The double dissociation in this version of the model (.212) is about half the size of that in the original model, indicating that semantic richness makes a significant contribution to the division of labor between syntax and semantics.

4.4. Summary of model manipulations

The most important observation from these manipulations is that the double dissociation between light and heavy verbs was maintained in each version of the model: light verbs were more disrupted than heavy verbs in the syntactically lesioned models, and heavy verbs were more disrupted than light verbs in the semantically lesioned models. The double dissociation was tested statistically, as the interaction between type of verb (light vs. heavy) and type of lesion (syntactic vs. semantic) in a 2×2 ANOVA, and found to be significant ($p < .001$) in each case. However, the size of the double dissociation varied, indicating that each variable played a role in determining the relative strengths of semantic and syntactic cues in the activation of light and heavy verbs. The greater frequency of occurrence of light verbs served to increase the light/heavy verb dissociation in semantically lesioned models and decrease it in syntactically lesioned models. The wider distribution of light verbs decreased the dissociation slightly in semantically lesioned models but increased it in syntactically lesioned models. The impoverished semantic representations of light verbs increased the dissociation with both types

Table 7
Input–output connection weights: Equal Features model

Input nodes	Output nodes										
	THE	BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP
DETERMINER	**3.43**	−1.66	−1.64	−1.63	−1.66	−1.67	−1.25	−1.27	−1.51	−1.18	−1.20
NOUN	−2.07	**1.45**	**1.41**	**1.45**	**1.41**	−1.60	−1.22	−1.28	−1.45	−1.13	−1.18
VERB	−2.05	−1.62	−1.56	−1.58	−1.62	**1.12**	**0.71**	**0.61**	**1.17**	**0.72**	**0.67**
MOTION	−0.21	−0.92	−0.88	−0.94	−0.86	**0.19**	**−0.01**	**−0.01**	−2.17	−1.70	−1.59
REMAIN	−0.22	−0.89	−0.91	−0.90	−0.90	−2.17	−1.71	−1.64	**0.19**	**0.04**	**0.04**
AIR	−0.18	−0.44	−0.47	−0.70	−0.71	−1.49	**0.49**	−0.74	−1.35	**0.47**	−0.77
PLAYFUL	−0.24	−0.75	−0.82	−0.41	−0.48	−1.41	−0.74	**0.46**	−1.40	−0.78	**0.47**
WING	−0.23	**−0.23**	**−0.22**	−1.47	−1.59	−0.90	−0.50	−1.22	−1.06	−0.54	−1.23
ANIMAL	−0.19	**1.37**	−1.66	−0.86	−0.79	−0.44	−0.24	−0.63	−0.53	−0.28	−0.59
DEVICE	−0.17	−1.59	**1.31**	−0.87	−0.76	−0.42	−0.25	−0.65	−0.52	−0.30	−0.64
PERSON	−0.27	−1.50	−1.58	**−0.21**	**−0.24**	−0.90	−1.22	−0.49	−0.94	−1.23	−0.44
MALE	−0.19	−0.80	−0.69	**1.41**	−1.61	−0.45	−0.58	−0.24	−0.49	−0.61	−0.33
FEMALE	−0.15	−0.87	−0.74	−1.62	**1.40**	−0.44	−0.60	−0.34	−0.50	−0.70	−0.23
NEUTRAL	−0.21	−0.59	−0.57	−0.57	−0.57	**0.76**	−1.45	−1.42	**0.80**	−1.39	−1.42

Note. Highlighted numbers indicate activated input features for each target output.

of lesion, because this factor contributes to a more radical division of labor for light and heavy verbs.

5. Study 4: single-word production

The single most common manifestation of aphasia is difficulty with word retrieval, a deficit most clearly evident in naming tasks. Picture naming entails a process of retrieving the concept represented by a visual image, mapping the conceptual representation onto a corresponding lexical representation, and retrieving the sound form of the lexical item. Semantically, this chain of events is similar to the stages of word-retrieval which occur during natural language production, but without the syntactic structure required for sentence production. Because of the different demands of the two tasks, aphasic patients sometimes show a difference in performance, even when the tasks use the same lexical items. In particular, agrammatic aphasics, who have difficulty processing syntactic structure, frequently have more difficulty producing words in sentences than words in isolation, whereas the word-finding abilities of anomic aphasics are more likely to benefit from the added cues that sentences provide (Pashek & Tompkins, 2002; Schwartz & Hodgson, 2002; Williams & Canter, 1982).

This difference in how agrammatics and anomics perform in naming as opposed to sentence production may be explained through the model. In our approach to production, the main difference between sentence production and naming is in the message. For a sentence, the message contains features for all elements of the sentence and these remain active throughout its production. For single-word naming, the message is simpler, corresponding only to the features of the word to be named. The assumption that the message is complete and static throughout production of a sentence requires that the syntactic-sequential states control serial order. The changing syntactic input serves as a kind of traffic cop, letting words of one type through at one time, and others at another time. Given this perspective, one might expect that syntax lesions would be associated with a much greater impairment in sentence production than in naming. To assess this claim, we used the models, trained for sentence production, to produce single words. In the sentence production simulation, the input for a given lexical item consisted of the syntactic and semantic features of that word, as well as the semantic features of the rest of the sentence. By contrast, the naming simulation involved presenting the model with only the input features specifying a given word (e.g., *NOUN*, *WING*, and *ANIMAL* for BIRD). Syntactically lesioned models were given only semantic features as input (e.g., *WING* and *ANIMAL* for BIRD), and semantically lesioned models received only syntactic features (e.g., *NOUN* for BIRD). As in sentence production, models with partial lesions were also tested.

The naming simulation examines whether the model, having learned its vocabulary in the context of connected speech, can reliably produce words in isolation. Table 8 shows the output activations and production probabilities for each noun and verb (the determiner was left out of the naming task) in the unlesioned, syntactically lesioned and semantically lesioned naming models. In the unlesioned model, both nouns and verbs are produced at high levels of accuracy. Given that the model was never trained on this task, its ability to name noun and verb targets represents a form of generalization to novel stimuli.

Table 8
Output activations and resulting production probabilities: naming task

BIRD	PLANE	BOY	GIRL	GO	FLY	SKIP	STAY	HOVER	HOP	Target	*p*
Unlesioned model											
0.889	0.322	0.243	0.252	0.074	0.075	0.037	0.064	0.092	0.042	bird	.991
0.309	**0.900**	0.232	0.234	0.070	0.082	0.035	0.067	0.088	0.040	plane	.993
0.241	0.233	**0.900**	0.316	0.071	0.028	0.087	0.075	0.035	0.099	boy	.993
0.242	0.233	0.321	**0.900**	0.073	0.030	0.090	0.075	0.038	0.088	girl	.993
0.061	0.057	0.058	0.067	**0.854**	0.516	0.500	0.305	0.192	0.202	go	.933
0.059	0.052	0.034	0.038	0.508	**0.784**	0.380	0.078	0.397	0.134	fly	.903
0.034	0.032	0.053	0.064	0.520	0.392	**0.754**	0.076	0.129	0.385	skip	.865
0.061	0.060	0.058	0.066	0.332	0.182	0.205	**0.845**	0.522	0.487	stay	.927
0.060	0.055	0.035	0.038	0.081	0.431	0.137	0.513	**0.751**	0.367	hover	.860
0.034	0.033	0.054	0.063	0.084	0.118	0.443	0.504	0.404	**0.701**	hop	.782
Syntactically lesioned model											
0.681	0.115	0.078	0.076	0.259	0.247	0.120	0.229	0.255	0.124	bird	.885
0.106	**0.711**	0.074	0.070	0.251	0.263	0.112	0.235	0.247	0.117	plane	.914
0.078	0.077	**0.704**	0.102	0.252	0.105	0.252	0.260	0.110	0.259	boy	.912
0.078	0.077	0.111	**0.687**	0.258	0.111	0.259	0.258	0.117	0.236	girl	.895
0.259	0.261	0.262	0.264	**0.623**	0.409	0.414	0.114	0.134	0.159	go	.676
0.253	0.242	0.171	0.167	0.226	**0.701**	0.302	0.024	0.299	0.103	fly	.894
0.159	0.159	0.247	0.256	0.234	0.294	**0.684**	0.023	0.088	0.318	skip	.865
0.261	0.270	0.265	0.261	0.123	0.126	0.154	**0.616**	0.415	0.415	stay	.648
0.255	0.251	0.173	0.164	0.024	0.329	0.101	0.236	**0.662**	0.302	hover	.837
0.161	0.166	0.249	0.253	0.025	0.080	0.359	0.230	0.305	**0.636**	hop	.776
Semantically lesioned model											
0.790	0.785	0.791	0.803	0.185	0.199	0.221	0.189	0.228	0.238	bird	.243
0.790	**0.785**	0.791	0.803	0.185	0.199	0.221	0.189	0.228	0.238	plane	.231
0.790	0.785	**0.791**	0.803	0.185	0.199	0.221	0.189	0.228	0.238	boy	.245
0.790	0.785	0.791	**0.803**	0.185	0.199	0.221	0.189	0.228	0.238	girl	.276
0.156	0.147	0.147	0.166	**0.780**	0.607	0.586	0.773	0.607	0.573	go	.390
0.156	0.147	0.147	0.166	0.780	**0.607**	0.586	0.773	0.607	0.573	fly	.069
0.156	0.147	0.147	0.166	0.780	0.607	**0.586**	0.773	0.607	0.573	skip	.056
0.156	0.147	0.147	0.166	0.780	0.607	0.586	**0.773**	0.607	0.573	stay	.363
0.156	0.147	0.147	0.166	0.780	0.607	0.586	0.773	**0.607**	0.573	hover	.069
0.156	0.147	0.147	0.166	0.780	0.607	0.586	0.773	0.607	**0.573**	hop	.049

Note. Highlighted numbers indicate activation levels of the targeted output nodes.

To compare the effects of the division of labor in the two tasks, refer back to Table 3, which shows the mean production probabilities in both the sentence production and naming simulations. In general, the naming results in the lesioned models roughly mirror those of the sentence production task: a double dissociation, reflecting a heavy verb advantage with syntactic lesions and a light verb advantage with semantic lesions, is observed in all but one of the versions of the model ($p < .001$). There were, however, some differences from the sentence production task. First, in the original model, the syntactic lesion results in significantly less impairment in both heavy and light verb naming: the mean probability of light verb naming is .662, compared to .208 in the sentence production task; the mean probability of heavy verb naming is .843, compared to .354 in sentence production. (Note that nouns are also much more accurately produced in single-word than in sentence production.) This better performance with naming reflects the fact that elements from within the sentence no longer compete with the verb for production. Such competition is quite damaging in sentence production when there is no syntactic input to direct activation to the appropriate word at the appropriate time. This result supports empirical findings that agrammatic aphasics are more accurate when naming words in isolation than in context (e.g., Schwartz & Hodgson, 2002; Williams & Canter, 1982). On the other hand, the absence of semantic context in naming is not noticeable in any of the semantically lesioned models, because all semantic connections are severed in both tasks.

The second task difference is in the relative size of the heavy/light dissociation in those syntactically lesioned models in which the range of distribution is artificially equalized across verb types. In the Equal Distribution model, a syntactic lesion gives rise to a heavy/light dissociation of .620 (compared to .097 in sentence production); in the Equal Frequency/Equal Distribution model a syntactic lesion results in a dissociation of .663 (compared to .185). The reason for these differential effects is that, in sentence production, equalizing the verbs' distributions allows light verbs to rely more on contextual semantic cues and therefore less on the syntactic-sequential cue. In naming, however, the contextual cues are lost, which necessitates that light verbs be produced from the remaining weakened cues. Therefore, the gulf between light and heavy verbs is widened, as indicated by whopping double dissociations of .951 in the Equal Distribution model and .804 in the Equal Frequency/Equal Distribution model.

The third task discrepancy is the size of the double dissociation in the Equal Features model. When light and heavy verbs have equal numbers of features, the heavy verb advantage of the syntactically lesioned model disappears in both tasks. Light verbs are more accurately produced because the frequency effect is now partially spread out across all its input nodes and is, therefore, not entirely wiped out by cutting off input from the syntactic *VERB* node. In sentence production, this makes heavy and light verbs approximately equivalent in accuracy. In naming, however, light verbs (with a mean probability of .933) actually perform better than heavy verbs (mean probability of .793), because the distributional cues for heavy verbs, which compensate for the light-verb frequency advantage in sentence production, are no longer present. Because light verbs are more accurately retrieved in *both* lesioned models, the double dissociation (.032) disappears in the naming task.

Naming performance was also examined with less extreme syntactic and semantic lesions. Fig. 3 shows how the production probabilities of nouns, light verbs, and heavy verbs vary as a function of the strength of input–output connections. As in the sentence production simulations, a gradual reversal of the light verb advantage is seen with syntactic lesions of increasing severity. However, as discussed above, the effects of these lesions are much less drastic than in the sentence production task, reflecting a relative performance advantage for agrammatics in single-word production tasks. Although no task difference was observed with complete semantic lesions, a slight performance advantage for sentence production over naming was revealed with partial lesions, because these preserve some of the contextual semantic input. With 50% lesions, for example, the probability of production in naming was .446 for heavy verbs, .802 for light verbs, and .878 for nouns, compared to .471, .877, and .928, respectively, in sentence production. This finding is also consistent with clinical and empirical findings that word-finding

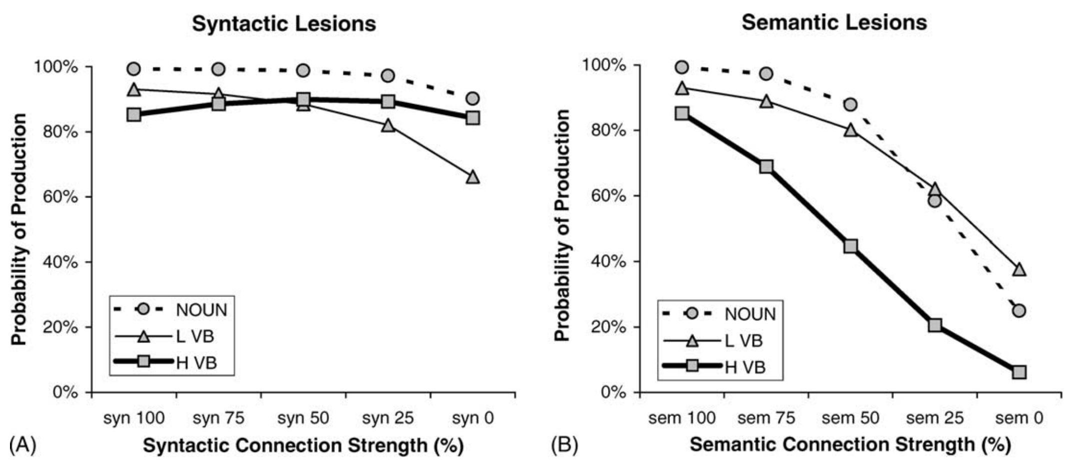

Fig. 3. Effect of lesions on probability of production: single-word production. To simulate partial lesions, connections between output nodes and either syntactic-sequential inputs (A) or semantic inputs (B) were systematically reduced to 75%, 50%, and 25% of their original strength. Unlesioned (100%) and completely lesioned (0%) models are also included for reference.

ability in anomic aphasics benefits from contextual support (e.g., Barton, Maruszewski, & Urrea, 1969; McCall, Cox, Shelton, & Weinrich, 1997; Pashek & Tompkins, 2002).

This dissociation in performance between the two tasks accords well with observations from the aphasia literature, as mentioned above. Furthermore, it may clarify the nature of the underlying sentence production deficit in agrammatism. What contributes to the agrammatic deficit in sentence production is not only the loss of syntactic markers, but also the interference from semantic features of the other lexical items within the sentence. When input from syntactic-sequential states is robust, this interference is largely eliminated, but without this information acting as traffic cop, features of the entire message are allowed to interfere with the target. Since the semantic features of words other than the single-word target are absent in the naming task, the agrammatic deficit is much less severe, solely reflecting the loss of syntactic-sequential cues on the retrieval of the target. The idea that agrammatism is related to both structural and lexical factors is not new (see for example, Rochon et al., 2000). What this model emphasizes, however, is that successful lexical retrieval depends not only on the characteristics of the targeted word, but also on its immediate competitors, as determined by the task. It remains to be seen how these factors would influence a more complex model of sentence production, but the comparison to naming appears to be as valuable in simulations as in experimental data.

6. General discussion

Although the model is an extremely simplified characterization of lexical access in sentence production, it provides insight into the nature of linguistic representations and their breakdown in aphasia. Most importantly, a double dissociation between heavy and light verbs can be simulated without postulating separate representations, provided that the model incorporates the following assumptions: that light verbs have fewer features than heavy verbs; that such features constitute a subset of heavy verb features; and that the contributions of semantics and syntax are determined by a learning algorithm that divides up the labor of activating output. The dissociation is most apparent when heavy verbs are less frequent and more constrained in their distribution, as is the case in natural language distributions. This combination of characteristics makes the syntax more important in the activation of light verbs. Their simpler, less specific semantics and wider co-occurrences with other sentence constituents make for relatively weak semantic cues, thus forcing the syntax to bear much of the responsibility for successful access. This division of labor enables the model to account for several aspects of language disruption in aphasia: the double dissociation between heavy and light verbs observed in agrammatic and anomic aphasics (Breedin et al., 1998); the more general dissociation between content and function words (e.g., Goodglass & Kaplan, 1983; Pulvermüller, 1995); and the dissociation between tasks involving single-word and sentence production in agrammatic aphasia (e.g., Schwartz & Hodgson, 2002). Furthermore, the division-of-labor hypothesis can accommodate aspects of aphasic performance which are problematic for other hypotheses—specifically, the fact that the relationship between underlying impairments and their behavioral manifestations is often not straightforward, and the tendency for certain characteristic deficits to cluster together in different types of aphasia.

6.1. The non-transparent relationship between deficits and underlying impairments

Because of the indirect relationship between the underlying impairment and its behavioral manifestations, the division-of-labor hypothesis provides a more powerful explanatory tool than transparent, modular models of impairment. It can account for double dissociations between linguistic structures for which there is no independent reason to assume a separation of representation or processing, like light and heavy verbs. The over-production of light nouns such as “thing” and “stuff”, another hallmark of anomia, might be assumed to reflect the same trade-off of semantic and syntactic cues.

Other non-transparent relationships have been proposed to account for dissociations in aphasic performance. For example, category-specific lexical access deficits in aphasia, such as animate versus inanimate things, objects versus tools, or nouns versus verbs, have been related to differences in the underlying semantic representations of different categories of words (e.g., Bird et al., 2000; Farah & McClelland, 1991; Shallice, 1988; Warrington & McCarthy, 1987; again, see Caramazza & Shelton, 1998 for further discussion). Farah and McClelland (1991) demonstrated that a frequently observed dissociation in naming living and non-living things can be simulated by selective damage to visual and functional attributes, respectively. Bird and colleagues (2000) further suggest that noun/verb dissociations may be related to a differential reliance on sensory and functional attributes. These hypotheses are attractive, not only because they are able to account parsimoniously for a variety of deficit patterns in aphasia, but because they are consistent with other neuropsychological evidence. In particular, the areas of the brain that are activated during retrieval of these different types of words have been found to dissociate: sensory association areas are implicated in the processing of living things and objects (especially animate), while areas anterior to Broca’s areas, close to motor cortex, are more strongly associated with the processing of tools and action words (Martin et al., 1995, 1996; Peterson et al., 1988; Tranel, Adolphs, Damasio, & Damasio, 2001).

A strictly modular account would require that types of structures which are behaviorally dissociated be represented separately in the lexicon and, thus, independently vulnerable to disruption (e.g., Tranel et al., 2001; Ullman et al., 1997). In most cases, this may be the most parsimonious explanation to account for observed dissociations. However, when there is no independent evidence to support separate representations—as is the case for light and heavy verbs—the modular account is on shaky ground. The division-of-labor hypothesis provides an alternative explanation which may be preferable in such cases. For the purposes of the present study, a division of labor between syntax and semantics provides a natural explanation which is motivated by normal language models and neuropsychological evidence.

6.2. *The co-occurrence of symptoms in aphasia*

In addition to explaining how functional dissociations can be indirectly related to underlying impairments, the distributed nature of processing in connectionist models also allows for multiple dissociations to arise from the same underlying cause. Given its representational and lesioning assumptions, the present model results in a division of labor that favors syntactic-sequential information for function words ("the") even more strongly than for light verbs. Similarly, Bastiaanse and colleagues found an association between the production of determiners and verbs in German agrammatic aphasics (Bastiaanse, Rispens, Ruigendijk, Rabadan, & Thompson, 2002). On the other hand, nouns pattern with heavy verbs, relying more on semantic than syntactic-sequential information. Although the network was not set up to simulate content/function word or noun/verb dissociations, the results suggest that such dissociations might have an underlying cause common to the light/heavy verb dissociation. The content/function word distinction is, by definition, a semantic/syntactic distinction which, in traditional language production models, is instantiated in different retrieval routes (e.g., Garrett, 1988). The division-of-labor concept provides an alternative, quantitative way of accounting for this dissociation within a connectionist framework. The noun/verb dissociation may have a similar basis: verbs have more complex grammatical representations (Bates et al., 1991; Breedin & Martin, 1996; Collina, Marangola, & Tabossi, 2001; Kim & Thompson, 2000), whereas nouns arguably have richer semantic representations (Bird et al., 2000; Caramazza & Shapiro, in press). The inconsistency of noun/verb dissociations in the literature may be related to the fact that their relative "heaviness", or semantic richness, has not been controlled. In addition to dissociations in the retrieval of different lexical types, the division of labor also shows potential in accounting for the task differences observed. Semantic context, to the extent that it is preserved, plays a facilitative role for anomic aphasics, but becomes a potential source of confusion when the syntactic-sequential cues are disrupted, as in agrammatism.

Underlying impairments along a syntactic-semantic dimension that is common to all these dissociations can naturally account for the frequently observed co-occurrence of these symptoms in many agrammatic and anomic patients. In modular theories, on the other hand, such symptom complexes can only be explained by coincidental damage to neurally proximal representations. For example, in the words of Miceli and colleagues (1984):

> [T]here appears to be no theoretically defensible connection between the omission of main verbs and the omission of grammatical markers in agrammatic speech: No theory of language production has

> been proposed which when appropriately "lesioned" (with a single lesion) predicts the omission of grammatical markers and verbs. The co-occurrence of these two symptoms, then, must be considered to be a consequence of an impairment to independent cognitive systems that tend to co-occur because of the neural proximity of the two systems . . . . (pp. 217–218)

Just such a functional association between symptoms is proposed here: the division-of-labor hypothesis predicts that words and tasks which rely more heavily on semantic representations will be disproportionately disrupted in anomia, whereas those which are more reliant on syntactic-sequential states will suffer more in agrammatism.

Of course, one of the hallmarks of aphasic syndromes is variability. Certain symptoms tend to co-occur, but may not in all patients. These dissociations may not be observed in all agrammatics and anomics (Jonkers & Bastiaanse, 1998), or may not constitute real double dissociations (Joanette & Goulet, 1991). In fact, Appelbaum and Bates (1999) point out that the incidence of double dissociations, as evidence for independence between functions, is overestimated because of the failure to consider correlations between measures. However, these observations are less problematic for connectionist explanations than they are for modular accounts. Because of the continuous and quantifiable nature of the dimensions assumed to underlie impairments in connectionist models (such as the relative connection strengths of syntactic and semantic inputs in the present study), continuity in the degree to which operations dissociate would likewise be expected. In modular models, deficits are frequently described in an all-or-none fashion, although more recent studies do attempt to account for variations in performance. For example, Ullman and colleagues (1997) make the proposal that relative impairments might arise as a consequence of a lesion that affects only part of a brain system. However, the graded nature of connectionist models allows a more motivated and straightforward account of the occurrence of partial and mixed lesions.

6.3. *The overactivation/underactivation dissociation*

There is one other unanticipated finding from this study which echoes an empirically observed dissociation in aphasia, and thus deserves mention. As illustrated in Table 1, syntactic lesions of the present model resulted in low levels of activation overall, whereas semantic lesions resulted in relatively high activation levels. Because the probability of production of the targeted output item depends on its activation level *relative* to the activation levels of all other possible outputs, these under- and over-activation patterns are not directly reflected in the accuracy levels. For example, when FLY is targeted in the syntactically lesioned model, its activation level is 0.456, and its probability of correct production is .460; when FLY is targeted in the semantically lesioned model, its activation level is higher, at 0.607, but its production probability is only .069, because the activation levels of competing words are also much higher.

Why does the model exhibit overactivation with semantic lesions and underactivation with syntactic lesions? The answer relates to the learning algorithm and the nature of the semantic and syntactic representations. Because the semantic representation of the entire message is on during the sentence, it follows that, at any given point in the sentence, most of the active features contribute more to the activation of non-target words than to the target. Under these conditions, inhibitory connections develop between the active features and the lexical items that they incorrectly turn on. With a semantic lesion, these inhibitory connections are eliminated, with the consequence that items are overactive. Consider, for example, the sentence "The bird flies" (see Table 2 for connection weights). At the time that BIRD is to be produced, the active semantic features are *ANIMAL*, *WING*, *MOTION*, and *AIR*. The summed input from these features to BIRD is inhibitory ($1.21 + -0.45 + -1.05 + -0.03 = -0.32$), and, of course, even more strongly inhibitory to BIRD's main competitor, PLANE ($-1.59 + -0.45 + -1.04 + -0.10 = -3.18$). It is the positive syntactic input from the *NOUN* node to each of the nouns (1.33 to BIRD, and similar values to the other nouns) that makes BIRD's total net input positive, leading to a high activation level and a high selection probability. If you take away the semantic contribution to net input, BIRD is even more active than under normal conditions. However, accuracy is extremely poor because all of the other nouns are equally active.

The syntactic represention is quite different. At any given time during sentence production, there is only one active syntactic-sequential feature (e.g., *NOUN*). Because this feature is predictive of the target intended to be produced at that time (e.g., BIRD), an excitatory connection is created between them. In the process of learning the training sentences, this occurs for each of the output lexical items. The syntactic-sequential feature thus controls output by sending positive input to all lexical items consistent with the syntactic-sequential state indicated by the feature. If this input is eliminated by a syntactic lesion, the target word and its competitors all become severely underactivated. In sum, our assumptions about a static featural representation for semantics and a relatively sparse dynamic representation of syntactic states, lead to different roles for inhibition and excitation in the two systems, differences that lead to over- and underactivation for anomic and agrammatic lesions, respectively.

This observation is intriguing because of its correspondence to a series of studies by Milberg, Blumstein and colleagues showing a dissociation between Broca's and Wernicke's aphasics in the degree of semantic priming shown relative to normal subjects (Milberg & Blumstein, 1981; Milberg, Blumstein, & Dworetsky, 1987, 1988a, 1988b). This difference was most clearly evident in a lexical decision task in which words were primed either by semantically related words (e.g., *cat–dog*), by phonological distortions of the semantic prime (e.g., *gat–dog*; *wat–dog*), or by unrelated words (e.g., *table–dog*). Normal subjects showed a monotonic relationship between the degree of phonological distortion of the prime (zero, one, or two phonetic features) and the amount of priming exhibited (Milberg et al., 1988a). Broca's aphasics showed no priming for either of the distorted primes, whereas Wernicke's aphasics showed as much priming for distorted primes as for the undistorted semantically related primes (Milberg et al., 1988b). These results were interpreted as illustrating an underactivation of the lexicon in Broca's aphasia and an overactivation of the lexicon in Wernicke's aphasia, with each deficit producing characteristic production patterns in these types of aphasia.

As an "existence proof" of the over/underactivation hypothesis, McNellis and Blumstein (2001) constructed a connectionist model of word recognition using the same priming conditions as in the behavioral experiments. The priming deficits of Broca's and Wernicke's aphasic patients were simulated by altering the resting activation levels of lexical items. Results of these simulations mirrored experimental results: the "Broca" model showed semantic facilitation of *dog* only in response to the semantically related prime, *cat*, whereas in the "Wernicke" model, semantic facilitation in response to the phonologically distorted primes was equivalent to the facilitation from *cat*. McNellis and Blumstein (2001) concluded that there is a critical range of resting activation in the normal lexicon. In Broca's aphasia, resting levels are below this range, with the result that incoming activation is insufficient for targets to reach the required threshold of activation for priming to occur. In Wernicke's aphasia, resting levels are elevated above the critical range, resulting in non-targeted items exceeding threshold.

Although the populations studied by Milberg, Blumstein and colleagues are not exactly equivalent to the agrammatic and anomic profiles of interest here, there is enough overlap in symptoms across the groups that it would not be unreasonable to hypothesize a similar dissociation in lexical activation levels in anomia and agrammatism. Our model naturally produces the hypothesized under- and over-activation but, in its present instantiation, these activation levels do not directly affect the accuracy of production. However, if the model had included a threshold of activation which items were required to surpass before being produced, then sentence production outcomes would have resulted in the frequent omission of items in the syntactically lesioned models. In the semantically lesioned models, several items would exceed the threshold, which, depending on the assumptions built into the model, might correspond behaviorally to the production of several paraphasic attempts at a target. Such a threshold can be conceived of in quasi-neural terms, in which the relevant units must be sufficiently active to exert an effect on connected units. Alternatively, it can be interpreted as an instantiation of the preserved ability of Broca's aphasics to monitor their language output, and thus inhibit potential errors, compared to an impairment in self-monitoring (corresponding to reduced thresholds) that is common in Wernicke's aphasics (Kolk & Heeschen, 1992). In any case, further simulations are required before firm conclusions can be drawn from the activation levels in the current model.

6.4. Conclusions and caveats

The ability of the model to demonstrate characteristics of language production that have been observed in aphasic subjects suggests that connectionist approaches have some promise in explaining dissociations in lexical access. Moreover, connectionist explanations have certain advantages over traditional, modular explanations. As discussed above, they can account for dissociations between items for which there is no *a priori* reason to assume representations in distinct modules; they provide a functional basis for the co-occurrence of symptoms observed in different types of aphasia, due to shared underlying processing components across items or tasks; and they allow for dissociations to vary in degree. As such, connectionist models provide a flexible and powerful alternative method of viewing language processing in the normal brain, and its breakdown secondary to brain damage.

Despite these potential advantages, results from the current simulation must be interpreted with a certain degree of caution, given the assumptions of the model. First and foremost, the model vastly simplifies the process of sentence production, yet it is assumed that the mechanisms of the simulation are representative of lexical access in spontaneous speech. As in any model, a degree of simplification was necessary in order to hold constant extraneous variables and reveal the operation of the particular variables of interest, in this case the relative roles of semantic and syntactic cues on the retrieval of heavy and light verbs. To pare down the process of sentence production, some arbitrary decisions were made, and some elements of the process were deliberately ignored. For example, all the nouns in the model were assigned two semantic features, like heavy verbs, although it is recognized that there is a continuum of "heaviness" for nouns as well. In addition, there are many aspects of the syntactic complexity of verbs that are not implemented in the model. The degree of complexity of the argument structures and thematic roles required by different verbs have been shown to contribute to the verb-retrieval deficits of agrammatic patients (e.g., Breedin & Martin, 1996; Thompson, Lange, Schneider, & Shapiro, 1997). Furthermore, it is generally recognized that verbs bear a great deal of syntactic responsibility in determining the structure of the rest of the sentence. There are often several structural alternatives for encoding a given message. Evidence suggests that the choice of syntactic frame may be influenced by, among other factors, the relative accessibility of the different lexical items that constitute these alternatives (Bock, 1982; Ferreira, 1996). To allow for this, an expanded model would need to represent alternative structures, and allow for the choice of structures to be influenced by activated lexical items and the message. In connectionist terms, the message and the lexical items would have learnable connections to the syntactic-sequential states.

By simplifying the process of sentence production, the model may also be seen to imply a simplistic view of the nature of agrammatic and anomic deficits. However, the model is by no means intended to capture the complexity of factors which contribute to the presentation of these aphasic syndromes. Rather, it is a means of representing one of the possible mechanisms by which symptoms characteristic of each syndrome can arise. The model's power, in fact, lies in its simplicity. The assumption that lexical information is divided into semantic and syntactic (or sequential) types of cues is independently motivated on linguistic and psychological grounds, and provides the basis for a number of dissociations in lexical access during production. Different categories of words naturally come to rely to varying degrees on semantic and syntactic cues, depending on their representational complexity, and their distribution and frequency in the language. What is critical to the model is that these cues "trade off" in importance—more reliable cues become stronger, while less reliable cues become weaker. Under normal circumstances, this division of labor provides an efficient and effective means of allocating resources during the production of language, and it can also account for some of the systematicities observed under conditions of disrupted language production. The model demonstrates one way in which these systematicities could emerge from learning and the consequent division of labor. More generally, we (along with many others, e.g., Chang, 2002; Christiansen & Chater, 2001; Plaut et al., 1996) encourage the development of explanations for language processing behavior that stem from the way in which language is learned.

Acknowledgments

This research was supported by a post-doctoral research fellowship (PD-1) to the first author from the Natural Sciences and Research Council of Canada (NSERC), and a research grant (DC-00191) to the second author from the National Institutes of Health (NIH). The authors are grateful to Eleanor Saffran, Myrna Schwartz, Susanne Gahl, David Plaut and two anonymous reviewers, for helpful feedback on earlier versions of the manuscript.

References

Appelbaum, M., & Bates, E. (1999). Quantifying dissociations in aphasia. *Brain and Language*, *69*(3), 313–316.

Barton, M., Maruszewski, M., & Urrea, D. (1969). Variation of stimulus context and its effect on word-finding ability in aphasics. *Cortex*, *5*, 351–365.

Bates, E., Chen, S., Tzeng, O., Li, P., & Opie, M. (1991). The noun-verb problem in Chinese aphasia. *Brain and Language*, *41*, 203–233.

Bencini, G., & Roland, D. (1996). *Verb access difficulties in agrammatic aphasic narratives*. Paper presented at the 70th Annual Meeting of the Linguistic Society of America, San Diego, CA.

Berg, T. (1988). *Die Abildung des Sprachproduktionprozess in einem Aktivationsflussmodell* [*A spreading activation model of speech production*]. Tuebingen, Germany: Max Niemeyer.

Berndt, R. S., Haendiges, A. N., Mitchum, C. C., & Sandson, J. (1997). Verb retrieval in aphasia. 2. Relationship to sentence processing. *Brain and Language*, *56*, 107–137.

Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia. 1. Characterizing single word impairments. *Brain and Language*, *56*, 68–106.

Bird, H., & Franklin, S. (1995/1996). Cinderella revisited: A comparison of fluent and non-fluent aphasic speech. *Journal of Neurolinguistics*, *9*(3), 187–206.

Bird, H., Howard, D., & Franklin, S. (2000). Why is a verb like an inanimate object? Grammatical category and semantic category deficits. *Brain and Language*, *72*, 246–309.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*(1), 1–47.

Bock, K., & Levelt, W. J. M. (1994). Language production. In M. A. Gernsbacher (Ed.), *The handbook of linguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bowles, N. L., & Poon, L. W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology*, *40*(1), 71–77.

Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 269–286). Cambridge, MA: MIT Press.

Breedin, S. D., & Martin, R. C. (1996). Patterns of verb impairment in aphasia: An analysis of four cases. *Cognitive Neuropsychology*, *13*(1), 51–91.

Breedin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: An effect of complexity. *Brain and Language*, *63*, 1–31.

Breen, K., & Warrington, E. K. (1994). A study of anomia: Evidence for a distinction between nominal and propositional language. *Cortex*, *30*, 231–245.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208.

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*, 788–790.

Caramazza, A., & Shapiro, K. (in press). Language categories in the brain: Evidence from aphasia. In L. Rizzi & A. Belletti (Eds.), *Structures and beyond*. Oxford, UK: Oxford University Press.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*(1), 1–34.

Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, *26*(5), 609–651.

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–229.

Christiansen, M. H., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CT: Ablex.

Collina, S., Marangola, P., & Tabossi, P. (2001). The role of argument structure in the production of nouns and verbs. *Neuropsychologia*, *39*(11), 1125–1137.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 318–344.

Damasio, A. R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences USA*, *90*, 4957–4960.

Danieli, A., Giustolisi, L., Silveri, M. C., Colosimo, C., & Gainotti, G. (1994). Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. *Neuropsychologia*, *32*(11), 1325–1341.

Dell, G. S. (1986). A spreading activation theory of retrieval in speech production. *Psychological Review*, *93*, 283–321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*(4), 313–349.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*(1), 123–147.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838.

Druks, J. (2002). Verbs and nouns—A review of the literature. *Journal of Neurolinguistics*, *15*, 289–315.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*(1), 71–99.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Farah, M., & McClelland, J. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*, 339–357.

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, *35*, 724–755.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27–52.

Gagnon, D. A., Schwartz, M. F., Martin, N., Dell, G. S., & Saffran, E. M. (1997). The origins of formal paraphasias in aphasics' picture naming. *Brain and Language*, *59*(3), 450–472.

Garrett, M. (1992). Disorders of lexical selection. *Cognition*, *42*, 143–180.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York, NY: Academic Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: Vol. III. Language: Psychological and biological aspects*. Cambridge, UK: Cambridge University Press.

Gleason, H. A. (1961). *An introduction to descriptive linguistics* (Rev. ed.). New York, NY: Holt, Rinehart and Winston.

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders*. Philadelphia, PA: Lea & Febiger.

Gordon, J. K. (1998). The fluency dimension in aphasia. *Aphasiology*, *12*(7/8), 673–688.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279.

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, *8*, 191–219.

Harm, M. W. (1998). *Division of labor in a computational model of visual word recognition*. Unpublished doctoral dissertation, Department of Computer Science, University of Southern California, Los Angeles, CA.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528.

Harm, M. W., & Seidenberg, M. S. (submitted). *Division of labor in a multicomponent connectionist model of reading. I. Computing the meanings of words*.

Hillis, A. E., & Caramazza, A. (1995). Representation of grammatical categories of words in the brain. *Journal of Cognitive Neuroscience*, *7*(3), 396–407.

Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbances. In R. Jakobson & M. Halle (Eds.), *Fundamentals of language*. The Hague: Mouton.

Jesperson, O. (1965). *A modern English grammar on historical principles*. London, UK: Allen & Unwin.

Joanette, Y., & Goulet, P. (1991). Text-level representations as one determinant for lexical retrieval and sentence production deficits in aphasia: Comments on L.B. Zingeser and R. Sloan Berndt "Retrieval of nouns and verbs in agrammatism and anomia". *Brain and Language*, *41*, 590–596.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences USA*, *96*, 7592–7597.

Jonkers, R., & Bastiaanse, R. (1998). How selective are selective word class deficits? Two case studies of action and object naming. *Aphasiology*, *12*(3), 245–256.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York, NY: Appleton-Century-Crofts.

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, *74*, 1–25.

Kegl, J. (1995). Levels of representation and units of access relevant to agrammatism. *Brain and Language*, *50*, 151–200.

Kohn, S. E., Lorch, M. P., & Pearson, D. M. (1989). Verb finding in aphasia. *Cortex*, *25*, 57–69.

Kolk, H., & Heeschen, C. (1992). Agrammatism, paragrammatism and the management of language. *Language and Cognitive Processes*, *7*(2), 89–129.

Lapointe, J. S. (1985). A theory of verb form use in the speech of agrammatic aphasics. *Brain and Language*, *24*, 100–155.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardini, A., & Pinna, G.-D. (2002). Verb-noun double dissociation in aphasic lexical impairments: The role of word frequency and imageability. *Brain and Language*, *81*, 432–444.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behaviors. *Psychological Review*, *89*, 483–506.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York, NY: Springer-Verlag.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*, 102–105.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.

McCall, D., Cox, D. M., Shelton, J. R., & Weinrich, M. (1997). The influence of syntactic and semantic information on picture-naming performance in aphasic patients. *Aphasiology*, *11*(6), 581–600.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

McNellis, M. G., & Blumstein, S. E. (2001). Self-organizing dynamics of lexical access in normals and aphasics. *Journal of Cognitive Neuroscience*, *13*(2), 151–170.

Menn, L., & Obler, L. (Eds.). (1990). *Agrammatic aphasia*. New York, NY: Academic Press.

Miceli, G., Silveri, M. C., Villa, G., & Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex*, *20*, 207–220.

Milberg, W., & Blumstein, S. (1981). Lexical decision and aphasia: Evidence for semantic processing. *Brain and Language*, *14*, 371–385.

Milberg, W., Blumstein, S., & Dworetsky, B. (1987). Processing of lexical ambiguities in aphasia. *Brain and Language*, *31*, 138–150.

Milberg, W., Blumstein, S., & Dworetsky, B. (1988a). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, *26*(4), 305–308.

Milberg, W., Blumstein, S., & Dworetsky, B. (1988b). Phonological processing and lexical access in aphasia. *Brain and Language*, *34*, 279–293.

Nicholas, M., Obler, L. K., Au, R., & Albert, M. L. (1996). *Brain and Language*, *54*(2), 184–195.

Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences on lexical retrieval in aphasia. *Aphasiology*, *16*(3), 261–286.

Peterson, S., Posner, M., Fox, P., Mintun, M., & Raichle, M. (1988). Positron emission tomographic studies of the cortical anatomy of single word processing. *Nature*, *331*, 585–589.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*, 73–193.

Plaut, D. (1995). Double dissociations without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*(2), 291–321.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500.

Plunkett, K., & Elman, J. L. (1997). *Exercises in rethinking innateness: A handbook for connectionist simulations*. Cambridge, MA: MIT Press.

Pulvermüller, F. (1995). Agrammatism: Behavioral description and neurobiological explanation. *Journal of Cognitive Neuroscience*, *7*(2), 165–181.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460–499.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning. II. Current research and theory* (pp. 64–99). New York, NY: Appleton.

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, *72*, 193–218.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*, 440–479.

Saffran, E., Schwartz, M., & Marin, O. S. M. (1980). Evidence from aphasia: Isolating the components of a production model. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk*. New York, NY: Academic Press.

Schade, U. (1992). *Konnektionismus-Zur Modellierung der Sprachproduktion* [*Connectionism: Modelling of language production*]. Opladen, Germany: Westdeutscher Verlag.

Schwartz, M. F. (1984). What the classical aphasia categories can't do for us, and why. *Brain and Language*, *21*, 3–8.

Schwartz, M. F., & Hodgson, C. (2002). A new multiword naming deficit: Evidence and interpretation. *Cognitive Neuropsychology*, *19*, 263–288.

Segalowitz, S. J., & Lane, K. C. (2000). Lexical access of function versus content words. *Brain and Language*, *75*, 376–389.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology*, *5*, 133–142.

Stemberger, J. P. (1985). *The lexicon in a model of language production*. New York, NY: Garland.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*(2), 135–170.

Tabor, W., & Tanenhaus, M. K. (1999). Dynamic models of sentence processing. *Cognitive Science*, *23*(4), 491–515.

Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*, *11*(4/5), 473–490.

Tranel, D., Adolphs, R., Damasio, H., & Damasio, A. R. (2001). A neural basis for the retrieval of words for actions. *Cognitive Neuropsychology*, *18*(7), 655–670.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, *9*(2), 266–276.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*(2), 105–143.

Warrington, E., & McCarthy, R. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, *110*, 1273–1296.

Widrow, B., & Hoff, M. E. (1988). Adaptive switching circuits. In J. A. Anderson & E. Rosenfeld (Eds.), *Neurocomputing: Foundations of research* (pp. 126–134). Cambridge, MA: MIT Press.

Williams, S. E., & Canter, G. J. (1982). The influence of situational context on naming performance in aphasic syndromes. *Brain and Language*, *17*, 92–106.

Williams, S. E., & Canter, G. J. (1987). Action-naming performance in four syndromes of aphasia. *Brain and Language*, *32*, 124–136.

Zingeser, L. B., & Berndt, R. (1988). Grammatical class and context effects in a case of pure anomia: Implications for models of language production. *Cognitive Neuropsychology*, *5*(4), 473–516.

Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, *39*(1), 14–32.