

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/265844166>

Inhibition in interactive activation models of linguistic selection and sequencing

Chapter · January 1994

CITATIONS

35

READS

301

2 authors, including:



[Padraig G. O'Seaghdha](#)

Lehigh University

34 PUBLICATIONS 1,372 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Form preparation and attention [View project](#)



Dynamic conceptual-lexical mapping [View project](#)

Inhibitory Processes in Attention, Memory, and Language

Edited by

Dale Dagenbach

*Department of Psychology
Wake Forest University
Winston-Salem, North Carolina*

Thomas H. Carr

*Department of Psychology
Michigan State University
East Lansing, Michigan*



ACADEMIC PRESS

A Division of Harcourt Brace & Company

San Diego New York Boston London Sydney Tokyo Toronto

This book is printed on acid-free paper. ∞

Copyright © 1994 by ACADEMIC PRESS, INC.

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Academic Press, Inc.

525 B Street, Suite 1900, San Diego, California 92101-4495

United Kingdom Edition published by

Academic Press Limited

24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in Publication Data

Inhibitory processes in attention, memory, and language / edited by
Dale Dagenbach, Thomas H. Carr

p. cm.

Includes bibliographical references and index.

ISBN 0-12-200410-8

1. Attention. 2. Memory. 3. Psycholinguistics. 4. Cognitive
neuroscience. I. Carr, Thomas H.

BF321.I54 1994

153.1 --dc20

93-31495

CIP

PRINTED IN THE UNITED STATES OF AMERICA

94 95 96 97 98 99 BB 9 8 7 6 5 4 3 2 1

Inhibition in Interactive Activation Models of Linguistic Selection and Sequencing

Gary S. Dell and Padraig G. O'Seaghdha

I. INTRODUCTION

The term *inhibition* is increasingly used in the recent literature on language production and comprehension. It is used in at least two distinct senses which we will label *molecular* and *molar*. In interactive activation models of language processing, inhibition is a molecular, quasi-neural process by which representational units send inhibitory signals to one another, or regulate their own activation by self-inhibition. In the experimental psycholinguistic literature, inhibition applies to a molar level of behavioral description that is often synonymous with interference. Molar inhibition may or may not entail inhibition at an underlying molecular level. For example, selection and decision processes involve competition, but they do not require that the competing elements inhibit one another. In this chapter, we focus on the use of the concept of inhibition in the interactive activation literature, but we will also examine the implications of our analysis for the interpretation of experimental data. We survey the wide range of functions served by inhibitory mechanisms in language processing and examine the question of whether inhibition is an identifiable process in language production and comprehension.

Inhibition, in the molecular sense, pertains to actual neural processes, or, in behavioral psychology, to hypothesized neural processes, such as lateral inhibition, that are seen in simple tasks involving the orientation of attention, the identification of simple stimuli, or the selection of a response. For basic processes such as these, there is a relatively close relation between psychology and neuroscience, or at least a prospect of integration between the two. In contrast, the study of language processes appears to be rather remote from the neural level. In the first place, language is not a basic process in cognition. It comprises several very different sorts of knowledge—phonological, syntactic, semantic, and pragmatic—and this knowledge is put to use in a variety of tasks: speaking, listening, reading, and writing. Moreover, even a relatively simple component of language processing, the retrieval of a word in production, is a complex multistep event in which many kinds of linguistic knowledge must be coordinated (e.g., Bock, 1982; Levelt et al., 1991). Given this complexity, it is not surprising that most psycholinguistic researchers have not shown much concern about whether a given phenomenon is, at core, based on excitation or inhibition. Of course, language processing models must have devices that select among alternatives or remove unwanted linguistic units, functions that suggest inhibition. But, in traditional psycholinguistic circles, the details of these devices have seemed secondary to questions regarding the nature of linguistic representations and the overall architecture of the processing system (e.g., Garrett, 1975; Fodor, Bever, & Garrett, 1974; Forster, 1976).

Another reason that inhibitory mechanisms in language processing have until recently not received much scrutiny is that the potential for bringing neural evidence to bear on putative inhibitory devices for linguistic functions has appeared slim. Although we can correlate language functions with brain areas through the study of aphasia and measures of brain metabolism, our ability to link such functions to particular inhibitory or excitatory neural processes is limited for obvious reasons. Experimental animals do not use language—and if they did, we would certainly be less inclined to use them as experimental animals.

Despite these considerations, two relatively recent developments in psycholinguistic theory have suggested the need for an examination of inhibitory mechanisms in language models. The first is the incursion of interactive activation, connectionist, neural-net, or parallel-distributed-processing models into word recognition (e.g., Elman & McClelland, 1986; Kawamoto & Zemblidge, 1992; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989) and other aspects of psycholinguistics (Bates & Elman, 1992; Dell, Juliano, & Govindjee, 1993; Elman, 1990; MacKay, 1987; MacWhinney & Leinbach, 1991; Plunkett & Marchman,

1991; Rumelhart & McClelland, 1986a; Stemberger, 1985; St. John & McClelland, 1990). Because these models make claims about behavioral data and compute by means of quasi-neural elements, including excitatory and inhibitory connections, they begin to provide a bridge between psycholinguistic phenomena and neural processes.

The second development in language research that requires a consideration of inhibitory processes is the explosion of research using priming techniques to study language comprehension and production. In a priming experiment, a subject either comprehends or produces a priming sentence, phrase, or word. At some point in this process, a target stimulus such as a written word is presented and the subject responds. For example, the subject might simply be required to read to word aloud as quickly as possible. The relation between the target and the linguistic prime is manipulated and the response time to the target indexes the processing of the sentence, phrase, or word. This method was originally used to study the activation of isolated words (e.g., Meyer & Schvaneveldt, 1971), was later extended to the study of lexically ambiguous words in context (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979), and has recently been applied to a wide variety of questions in both language production and comprehension (for reviews, see Dell & O'Seaghdha, 1992; McKoon & Ratcliff, 1992; Kawamoto, 1993). Because the effects of a related prime on the target are sometimes facilitatory (e.g., Meyer & Schvaneveldt, 1971; Tanenhaus et al., 1979), sometimes inhibitory (e.g., Blaxton & Neely, 1983; Colombo, 1986; Levelt et al., 1991), and sometimes both (e.g., Levelt et al., 1991; Schriefers, Meyer, & Levelt, 1990), a theory of inhibitory and excitatory mechanisms is required in order to interpret the data. In particular, we will consider how strongly the finding of behavioral inhibition between a prime item A and a target item B suggests a theory in which there is an explicit inhibitory connection from A to B. We will return to a consideration of the priming data after we examine the role of inhibition in interactive activation models.

Although most interactive activation models employ a combination of excitation and inhibition, we focus in this chapter on the role of inhibitory mechanisms in fulfilling psycholinguistic functions. In addition, we will emphasize language production, although issues in comprehension and word recognition will frequently arise.

II. INTERACTIVE ACTIVATION

Before we turn to psycholinguistic issues, we briefly outline the nature of interactive activation models. An interactive activation model is

a specific kind of connectionist or neural network model, the classic examples being Grossberg's (1978) adaptive resonance theory, and McClelland and Rumelhart's (1981) word recognition model. As McClelland (1987) describes the prototype of this type of model, computation takes place in a network of excitatory and inhibitory connections among simple units. These units represent hypotheses about input in the case of a perceptual system, and they represent potential actions in the case of a production system. Each unit has an activation level that is a function of the excitatory and inhibitory inputs it receives, and is often also influenced by its own recent activation level. The units are arranged in layers, most commonly in an abstraction hierarchy, with units representing more peripheral information occupying lower layers. For example, in a lexical network, one might find phonetic or articulatory features at the lowest level, followed by phonological segments (phonemes) at the next level, combinations of segments such as syllables and words at the next level, and semantic features or concept nodes at the highest level. The network computes by activating an outside layer, the lowest one in the case of perception and the highest in the case of production, and activation spreads throughout the network. The units typically connect only to other units in the same and adjacent layers. Connections within a layer are usually inhibitory, and those between layers are usually excitatory. Processing is often assumed to be interactive, in the sense that later levels can influence earlier levels. This occurs because connections run in both directions between levels, that is, there are both top-down and bottom-up connections.

It is important to keep in mind that not all interactive activation models have all of the preceding features. Rather, the models share a family resemblance, most exhibiting many of these characteristics and some manifesting additional features that we have not listed.

III. FUNCTIONS OF INHIBITION IN INTERACTIVE ACTIVATION MODELS OF LANGUAGE PRODUCTION

Saying a sentence involves making a great many decisions—decisions about lexical items (Should I say *dime* or *coin*?), about sentence structure (Should the active or passive be used? Which noun phrases should be expressed as pronouns?), about optional arguments (Do I need to say *with a bat* when I say *The girl hit the ball*?), about constituent order not determined by sentence structure (Is it *boys and girls* or *girls and boys*?), about intonational phrasing, pausing, and many other features. The informational sources and principles that guide these decisions are the primary objects of study in the field of language production (Levelt, 1989). Recently, models of production have adopted activation-

based approaches to decision making (e.g. Berg, 1988; Berg & Schade, 1992; Bock, 1982, 1986b; Bock, Loebell, & Morey, 1992; Dell, 1986, 1988; Dell et al., 1993; Dell & Reich, 1981; Eikmeyer & Schade, 1991; Harley, 1984, 1990; Houghton, 1990; Levelt, 1989; Levelt et al., 1991; MacKay, 1982, 1987; Martin, Weisberg, & Saffran, 1989; Martin & Saffran, 1992; Meyer & Gordon, 1985; Meyer, 1991; O'Seaghdha, Dell, Peterson, & Juliano, 1992; Peterson, Dell, & O'Seaghdha, 1989; Roelofs, 1992; Schade, 1992; Stemmer, 1985, 1990, 1991). In these models, constraints on production decisions are translated into excitatory and inhibitory inputs to units representing the options open to the system. The inputs modulate the activation level, or the probability of being active, of these units. Thus, decisions are, to a large extent, a function of the inputs to units as determined by the structure of the network.

The activation-based models of language production must make two distinct kinds of decisions: selection decisions and sequence decisions. Using standard linguistic terminology (cf. Eikmeyer and Schade, 1991), we label these paradigmatic decisions and syntagmatic decisions, respectively. *Paradigmatic decisions* involve selection for a particular role from a set of candidate units in the relevant linguistic category. For example, the system may need to choose *dog* rather than *cat*, or *eat* instead of *drink* when selecting lexical items, or /k/ instead of /d/ when choosing, say, the initial consonant of the syllable /kæt/. Another example is a choice of syntactic structure, such as a double object construction (*Mary gave John the ball*) versus a prepositional dative (*Mary gave the ball to John*).

In contrast to paradigmatic decisions, *syntagmatic decisions* involve the selection of the correct elements of sequences at the right time. When producing *big dogs*, the activation of *big* must initially dominate that of *dogs*, and then the relative activations must quickly flip-flop. Likewise, in the production of a particular word, for example *dog*, /d/ must initially dominate, then /ɔ/, and then /g/.

Both paradigmatic and syntagmatic decisions often depend on inhibitory processes in activation-based language models. In both cases, the correct unit must prevail over its competitors. Many models employ a combination of excitation and inhibition to achieve the desired patterns of activation, but because the nature of the competition among units is different in the paradigmatic and syntagmatic cases, the configurations of excitation and inhibition in the two cases must also differ. There are two principal differences between paradigmatic and syntagmatic decisions. First, for paradigmatic decisions, the relation between correct and competing units can be partly specified in advance, and to this extent is independent of the intended utterance. Units in the same linguistic categories are competitors. At the syntactic level, nouns compete with other nouns. At a phonological level, vowels compete with vowels.

Although the categorical structure may be complex, including subcategories and multiple membership, the structure of language tells us, beforehand, which items have the potential to incorrectly replace other items. We can see this paradigmatic principle at work in noncontextual speech errors, slips of the tongue in which a linguistic unit from outside the intended utterance replaces an intended unit. Most of the time, the replacing unit is a member of the same linguistically defined category as the replaced one (Fromkin, 1971; Garrett, 1975; MacKay, 1972).

In contrast, for syntagmatic decisions, the set of competitors depends not on categorical structure but on the content of the intended utterance. We subdivide syntagmatic decisions into two classes: noncreative and creative. In noncreative syntagmatic decisions, those involving a well-known sequence, for example, an idiom or the sounds that make up a familiar word, the competing elements are already stored in memory. In the case of saying the word *cat*, the segments /k/, /æ/, and /t/ are in competition, and the structure of the word determines how that competition should evolve over time. However, in the case of creative language use, such as the construction of novel sentences or phrases, the syntagmatic competition pattern is not so easily specified (see MacKay's, 1987, discussion). The same system that can store the ordering relations in the familiar phrase *Venetian blind*, must also be capable of producing and understanding the novel phrase *blind Venetian*.

The second critical difference between syntagmatic and paradigmatic decisions is that in the syntagmatic case, the problem is not what items to output but rather one of outputting the items in the correct serial order. Each competitor has to win the competition at some point. For example, although the /ɔ/ in *dog* must lose to /d/ initially, it must quickly become the winner of the next competition. Unlike the paradigmatic case, in which the unwanted competitors can be unceremoniously dispatched, rejected items must often be immediately resuscitated in syntagmatic processing.

Both of these differences suggest that syntagmatic decisions are more complex than paradigmatic ones. As we shall see, syntagmatic decisions, particularly creative syntagmatic decisions, require complex excitatory and inhibitory systems in activation-based models of production. We therefore consider the simpler paradigmatic case first.

IV. PARADIGMATIC CHOICES AND LATERAL INHIBITION

Interactive activation models of production and comprehension have, for the most part, used some kind of lateral inhibition in paradigmatic decisions. Units that represent mutually exclusive options send inhibition to one another with the result that the single option with the

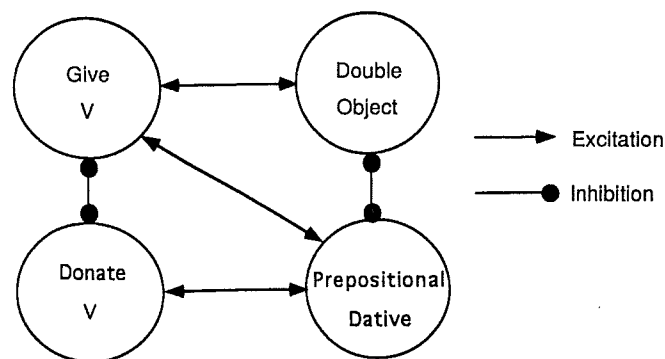


FIGURE 1

Lateral inhibition among competing verbs and competing syntactic structures, and excitation between compatible verbs and structures allows for lexical-syntactic interaction.

strongest support is eventually selected. This configuration is a potent decision-making system, one that the brain clearly adopts (Feldman & Ballard, 1982; Grossberg, 1978; Ratliff & Hartline, 1959). Interactive activation models of language production that use lateral inhibition among linguistic units at the same level include Berg (1988), Eikmeyer and Schade (1991), Harley (1984, 1990), Houghton (1990), Schade (1992), and Stemberger (1985, 1990, 1991). The decision to incorporate inhibition in these models was largely inspired by the success of visual and auditory word-recognition models (e.g., Elman & McClelland, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) and typing models (e.g., Rumelhart & Norman, 1982), all of which made extensive use of within-level inhibition. Specifically, these models proposed that nodes for sublexical (letters, phonemes) and lexical units inhibited others at the same level.¹

To illustrate lateral inhibition in language production, Figure 1 shows a small network of units representing two verbs, *give* and *donate*, and two syntactic options coded as single nodes, one for a double-object construction and one for a prepositional dative. The configuration embodies the approach to lexical-syntactic interactions present in the work of Bock (1982) and Stemberger (1985). Such a network will settle to a state in which a single verb and a single syntactic structure that is appropriate for that verb are active. The network uses lateral inhibition between the verbs and between the syntactic options along with

¹Because nodes for letters in the McClelland and Rumelhart visual model were duplicated for each serial position within a word, and all nodes in the Elman and McClelland auditory model were duplicated for each temporal position or "time slice," lateral inhibition only occurred between nodes that dealt with the same spatial or temporal positions. That is, the inhibition was used to resolve paradigmatic competition within a given position.

excitatory connections that specify the syntactic subcategorizations of the verbs. Note that an alternative network might have each verb inhibiting the syntactic structures that are not appropriate for it, rather than exciting the appropriate ones. For this problem, the excitatory solution would, in a complete model, use fewer connections, because there are fewer appropriate than inappropriate structures. The excitatory solution is also implicitly assumed in psycholinguistic theories that emphasize automatic activation of subcategorization (e.g., Boland, Tanenhaus, & Garnsey, 1990), but, aside from the number of connections, there is no compelling computational reason to prefer one over the other. A third possible configuration would have both excitatory connections from each verb to its appropriate structures and inhibitory connections to inappropriate structures, combined with lateral inhibition at each level.

Although many activation-based models have used lateral inhibition to aid decision making, several models of production use only excitation. However, when lateral inhibition is not present, there are other mechanisms for enhancing the activation of intended units, or otherwise setting them apart from competing units. For example, MacKay (1982) defined priming as a process by which nodes representing options within a given linguistic category (e.g., verb) gradually build up their activation levels. Eventually, one node reaches a threshold and becomes "activated." When a node is activated, its activation level becomes much greater and is sustained over a period of time. Thus, in MacKay's model, the decision mechanism is an absolute threshold and does not use lateral inhibition, but it achieves the desired result: a distancing of the winner of the race from the competition. Likewise, in Dell's (1986, 1988) language production model, the most highly activated node within a category is "selected" and linked to a particular slot in the production plan. Because linkage eventually entails a large boost of activation, the effect is somewhat similar to what obtains in MacKay's model when the threshold is reached. Roelof's (1992) model of lexical retrieval in production incorporates two selection mechanisms: a critical difference threshold, in which a target word node becomes available for selection when its activation is greater than that of all of its competitors by a certain amount, and a stochastic mechanism, in which the probability that the available target is selected is a function of the ratio of its activation level to the total activation of all competitors (see Luce, 1959). Again, the effect is similar to lateral inhibition, in that the most activated word ends up being selected.

A feature of lateral inhibition in some of the production models is that it applies within all pairs of units at a linguistic level (e.g., Eikmeyer & Schade, 1991; Harley, 1990; Stemmer, 1985). So all words inhibit all other words, all segment nodes inhibit all other segment nodes, and

so on.² Earlier however, we suggested that paradigmatic decisions might be profitably restricted, perhaps by linguistic category. For example, vowels would inhibit other vowels, but not consonants. In addition, paradigmatic competition might be further narrowed if the extent of lateral inhibition were related to similarity, with inhibition reserved for or more potent among similar units. This idea is implicit in the following literature on retrieval inhibition, in which tokens from the same semantic category appear to inhibit one another under certain conditions (e.g., Brown, 1981; Blaxton & Neely, 1983; Wheeldon, 1989). Likewise, Meyer and Gordon (1985) and MacKay (1970) have suggested that lateral inhibition occurs primarily between speech sounds that share features. Finally, the idea of similarity-based competition is inherent in many models of auditory and visual word recognition (see, e.g., Colombo, 1986; Forster & Davis, 1991; Marslen-Wilson & Tyler, 1980).

A. Similarity-Based Inhibition

The question of the role of similarity or neighborhood in determining the strength of lateral inhibition can be illuminated by considering the function of such inhibition and by examining how inhibitory connections might be acquired. A paradigmatic decision is, by definition, one in which a single unit is chosen from a predetermined set of potential competitors for a linguistic role; for example, the role of subject noun of a sentence. If only the members of the appropriate set, say nouns, are highly activated at the time the subject noun is being chosen, then lateral inhibition might be limited to that set. But even within the category of nouns, it makes sense that some nouns would be more potent competitors for a given role than others. In attempting to find the appropriate name for a particular aquatic mammal (*dolphin*), the noun *pencil* should not be much competition. It might, therefore, be useful to have *dolphin* inhibit semantically similar nouns such as *porpoise* to a greater extent than *pencil*.

On first consideration, such "smart" inhibition might seem unlikely. How could the semantic system allow inhibition among closely related competitors while retaining the ability to activate relevant information in a sufficiently wide-ranging manner? However, examination of the operation of error-correcting learning procedures suggests that such selective inhibition could indeed be learned. Similarity-dependent lateral

²Part of the reason that these models have such wide-ranging lateral inhibition is that it plays a role in their treatment of serial order as well as in paradigmatic decision making. Because any word has the potential to appear in a sentence with any other word, all words may have to inhibit one another if lateral inhibition is an important mechanism in creative syntagmatic decisions. See our discussion of the Eikmeyer and Schade model.

inhibition would result from the application of error-based algorithms (e.g., back-propagation; Rumelhart, Hinton, & Williams, 1986) in an interactive-activation architecture. These procedures change connection weights in response to feedback regarding target activation levels that is delivered while a model is being trained to produce a particular input-output mapping.

Consider the task of mapping from a semantic unit or set of units representing a particular concept to a single unit representing the corresponding word or lemma (e.g., Levelt, 1989; Roelofs, 1992). Further, let us suppose that there are potential connections from the semantic units to the word units, and potential lateral connections among word units, and that the excitatory or inhibitory character of these connections is set by the learning process. Under such conditions, lateral inhibition will develop between semantic neighbors roughly in proportion to their relatedness.

To see how this happens, consider the case of two word units, *dog* and *cat*. The word *dog* is connected to the features *canine* and *pet*, and *cat* to the features *feline* and *pet*. Assume that there are top-down excitatory connections from the semantic unit to the appropriate word units, such that the correct word is more activated by its two features than the other word. But, as of yet, there is no lateral inhibition. Now, when *pet* and *canine* are input to the network, both *dog* and *cat* are activated, let us say, $dog = .75$ and $cat = .25$, on a range of $[0,1]$. The target activation levels, provided by external feedback, however, are 1 for *dog* and 0 for *cat*. That is, the network wants *dog* fully active and *cat* completely inactive. Hence, there would be an error signal of .25 (target activation - actual activation) on *dog* and a signal of $-.25$ on *cat*. The weight change from unit *i* to unit *j* with error-based learning algorithms is in proportion to the activation of unit *i* and the error signal on unit *j*. So, some lateral inhibition ($.75 \times -.25 = -.1875$) would grow from *dog* to *cat*.³ In addition, a smaller amount of lateral excitation ($.25 \times .25 = .0625$) would grow from *cat* to *dog*. However, when *cat* is the intended word, the reverse effects occur (large growth of inhibition from *cat* to *dog*, with a smaller growth of excitation from *dog* to *cat*). The net result over the long run is that the two words inhibit one another.

B. Modularity or Interdependence of Perception and Production

If lateral inhibition is a function of similarity and error correction, the extent to which linguistic units inhibit should be related both to

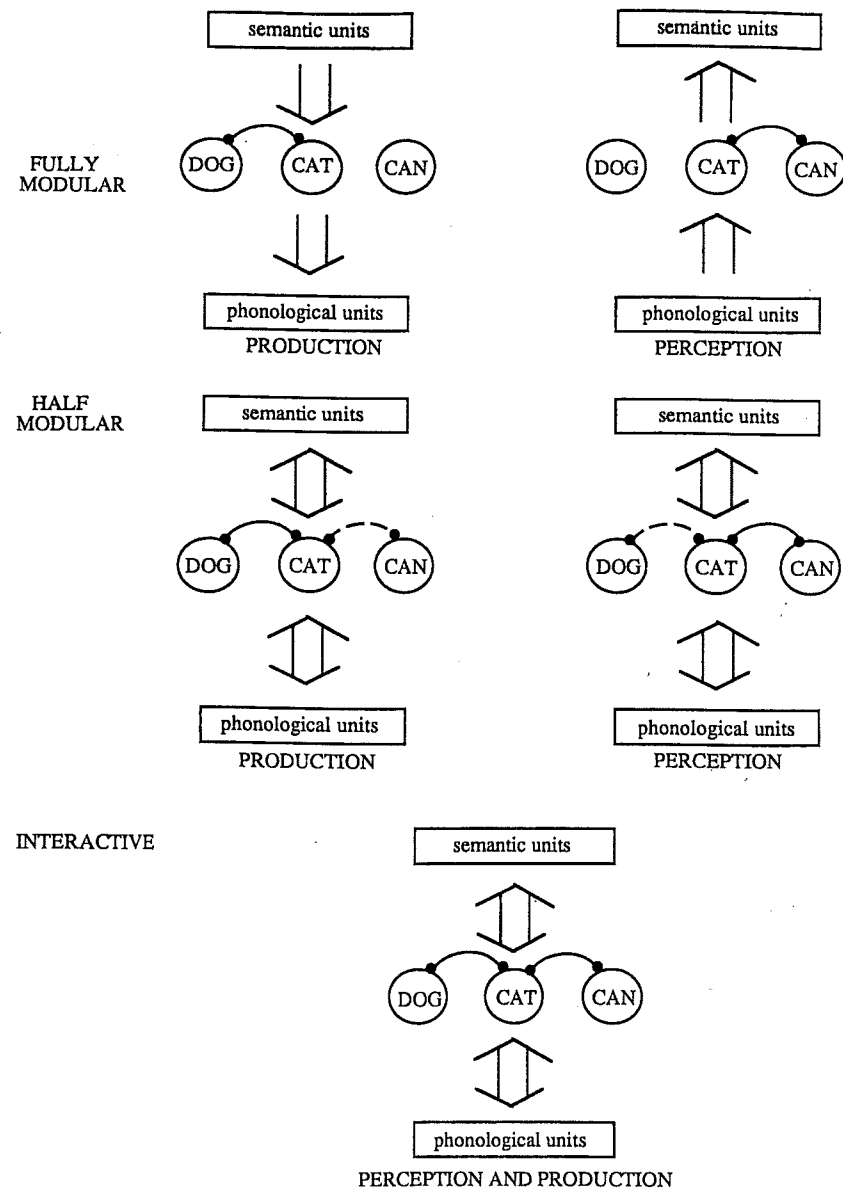
³This assumes a learning constant of 1, and an activation function in which activation of a node equals its net input.

similarity and to the linguistic task being performed. We have just seen how semantically related words might come to inhibit one another in the task of mapping from concepts to words. When the task involves word perception, either visual or auditory, the same considerations would lead to formally similar, instead of semantically similar words inhibiting each other. Hence, in production, *dog* and *cat* would mutually inhibit, whereas in perception, *cat* and *can* would inhibit one another. This logic suggests a number of possible configurations when the full architecture of the mental lexicon is considered. Figure 2 presents three of the possibilities.

The first possibility, which we call a *fully modular* architecture, has two characteristics: There are separate units and connections for perception and production, and there is no feedback or reverse flow of activation. In such a system, what we call *task-dictated lateral inhibition* would grow from an error-correction learning procedure: The *dog*-production and *cat*-production units would inhibit one another, as would the *cat*-perception and *can*-perception units. Because of the separateness of production and perception, and the fact that activation only flows in the task-dictated direction, each task's network does not experience the similarity relations governing competition in the other task.

In contrast, a *half-modular* architecture retains the separate production and perception networks, but allows a two-way flow of activation in each. In this case, the task-dictated pattern of similarity-based lateral inhibition would occur as before, but there would also be a *reverse* pattern pertaining to the other task—the *cat* and *can* production units would inhibit, and the *dog* and *cat* perception units would inhibit. This reverse pattern would arise from feedforward and feedback within each network. During production, when the target word is *cat*, not only *dog* but also *can* is a competitor because of the flow of activation from *cat* to its phonological units (*/k/*, */æ/*, and */t/*) and back to formally related words. Likewise, during perception, the target *cat* activates semantic units that then activate *dog*. The result, in both perception and production, is that a target word encounters both semantic and form-related competitors, and needs to inhibit them. However, with the half-modular architecture, the task-dependent lateral inhibition connections should be much stronger than the reverse ones because the task-dictated competitors (*dog* for *cat* in production, *can* for *cat* in perception) would be activated earlier and to a greater extent than the competitors activated by feedback (*can* for *cat* in production, and *dog* for *cat* in perception).

The third possibility we consider is a *fully interactive* architecture (see Figure 2). This configuration involves a single network for production and perception and consequently must allow activation to flow in both directions (e.g., MacKay, 1987). In this situation, a given word unit would develop lateral inhibition to both semantic competitors (from

**FIGURE 2**

Three architectures and their associated patterns of lateral inhibition, assuming that lateral inhibition reflects experienced competition.

production and, to a smaller extent, from feedback during perception) and formally related competitors (primarily from perception, and secondarily from production). So, the fully interactive architecture resembles the half-modular one in that an error-correction learning rule creates the reverse pattern of lateral inhibition as well as the task-dictated pattern. In the fully interactive case, however, the reverse pattern would be as strong as the task-dictated one because the same nodes are used in both tasks. Any tendency for *cat* to inhibit *dog* in production should also be observed in perception.

C. Empirical Status of Paradigmatic Inhibition

We now turn to an evaluation of the experimental evidence for paradigmatic lateral inhibition. As we indicated earlier, priming procedures provide the major source of data on facilitation and inhibition in language processing. Therefore, this is where we will look for evidence concerning the prevalence of inhibition.

Several obstacles stand in the way of a straightforward evaluation of the evidence concerning lateral inhibition. Chief among them is the fact we alluded to earlier—that we do not have direct access to inhibitory mechanisms in language processing. Further, if inhibition exists, it may well be obscured by concurrent excitatory activation. To take one example, the fully interactive configuration we have just considered suggests the possibility that lateral inhibition exists between semantically related words in perception. Behaviorally, it is well known that *cat* facilitates *dog* in perception, but is this effect purely excitatory, or is it possible that semantic facilitation merely dominates lateral inhibition? Another obstacle to the evaluation of inhibition in priming data is that experimental tasks are not purely perceptual or productive. Rather, they involve some mixture of these processes. For example, responses in perceptual priming experiments are productive, and eliciting displays in production experiments necessarily entail perception. As we shall see, the mapping between tasks and levels of processing is very difficult to specify.

Keeping these concerns in mind, our strategy will be to first look for inhibition in the most clearcut cases, the task-dictated cases for which all configurations in Figure 2 predict lateral inhibition. Therefore, we will look for evidence of semantic inhibition in production and for form-related inhibition in perception. Interestingly, both the possibility of semantic inhibition in production and of form-related inhibition in perception have emerged in the literature as counterintuitive phenomena, which goes against the grain of facilitation-based theories of similarity effects. Where we do find evidence of task-dictated inhibition, we

will consider whether processes other than lateral inhibition might account for it. Our analysis suggests that if we do not find lateral inhibition in the task-dictated cases, we should not find it elsewhere. We will keep this in mind in evaluating the status of reverse effects—semantic inhibition in perception and form-based inhibition in production.

1. Semantic Inhibition in Production There is now considerable evidence that inhibition may occur between related concepts in the process of semantic retrieval. The experimental literature on this topic may be traced at least as far back as the category-instance experiments of Loftus and associates (e.g., Loftus, 1973; Loftus & Loftus, 1974). These experiments measured latency to retrieve the appropriate instance (e.g., APPLE) given a category label and a letter restrictor (e.g., FRUIT-A). Though these early studies found that the second retrieval from the same category (e.g., FRUIT-A; FRUIT-P vs. VEHICLE-C; FRUIT-P) was facilitated, later studies found that repeated retrievals from the same category led to inhibition (e.g., Blaxton & Neely, 1983; Bowles & Poon, 1985; Brown, 1979, 1981; Roediger, Neely, & Blaxton, 1983; Wheeldon, 1989; and see Roediger & Neely, 1982, for a review of other retrieval inhibition effects). Blaxton and Neely (1983) found inhibitory effects when subjects generated the names of four category instances prior to a critical retrieval trial, but not when they read them. This suggests that prelexical retrieval processes must be engaged for inhibition to occur. In fact, this rather than the number of retrievals may be the critical factor. Studies using definition tasks that involve difficult retrieval show inhibition effects following a single retrieval (see Bowles & Poon, 1985; Wheeldon, 1989).

Although the evidence suggests that retrieval inhibition in production is a real phenomenon, the nature of the effect is not evident. First, the evidence that prelexical retrieval is involved suggests that the effect may not be due to inhibition among lexical units. We explore this question further in the subsection on reverse effects. Second, the molar inhibition is not necessarily mediated by molecular lateral inhibition. Earlier, we contrasted models of language production with lateral inhibition (e.g., Eikmeyer & Schade, 1991; Stemmer, 1985) and models that eschew it (Dell, 1986; MacKay, 1987; Roelofs, 1992). In the latter models, the activation of conceptually related items creates a situation in which they may compete for selection. The models need only assume that the decision process is more difficult when more than one exemplar is highly activated to account for retrieval inhibition. We note that models that incorporate a differential selection criterion (e.g., Roelofs, 1992) are better equipped for this purpose.

In fact, Roelofs (1992) provides one clear case where molar semantic inhibition in production likely is *not* solely due to molecular lateral

inhibition. Roelofs found that picture naming was inhibited by a superimposed semantically related word. For example, producing *dog* to a picture of a dog was slowed by the word *fish* printed on the picture. This result cannot be due to hardwired inhibition between pairs of animal words, however, because the inhibitory effect is present only when the interfering word is known to be a potential response in the task. When *fish* is *not* a potential response, semantically related words facilitate picture naming. Thus, in this case, the competitive relations are dynamic rather than hardwired. We suggest that this characterization may apply generally to semantic inhibition. Similarity-based behavioral inhibition in production may be a real phenomenon, but we think it unlikely that the effect is mediated by fixed inhibitory connections among semantically related words.

2. Form-Related Inhibition in Word Recognition We turn now to the possibility of lateral inhibition among formally similar words in word recognition. The notion of lateral inhibition is used in interactive activation models of visual and auditory word recognition, including those listed earlier (e.g., McClelland & Rumelhart, 1981; Elman & McClelland, 1986). In addition, the potential for similarity-based competition of the kind we are considering is directly related to the distributional properties of lexical neighborhoods (e.g., Andrews, 1989; Coltheart, Davelaar, Jonasson, & Besner, 1977; Glushko, 1979; Luce, Pisoni, & Goldinger, 1990; Marslen-Wilson & Tyler, 1980; and see Forster & Davis, 1991; Segui & Grainger, 1990; Slowiaczek & Hamburger, 1992; Taraban & McClelland, 1987). Although the models proposed in these references differ in many important respects, all of them conceive of auditory or visual word recognition as a discrimination problem. From the point of view of our analysis of similarity, such discrimination is more difficult in the more highly populated lexical neighborhoods that are inhabited by many formally similar words. Therefore, independent of other considerations, the potential exists for the growth of similarity-based lateral inhibition as a function of neighborhood density.

Although we have identified a basis for the existence of similarity-based lateral inhibition, it is difficult to assess the status of neighborhood density effects for some of the reasons outlined earlier. In particular, interactive activation models accommodate both lexical facilitation due to feedback from sublexical representations and inhibition due to lexical competition, and these effects may be modulated by other lexical characteristics such as frequency. Therefore, it is difficult to isolate inhibitory and facilitatory effects of neighborhood structure. In any case, the empirical picture on this issue is incomplete (but see Andrews, 1989; Forster, Davis, Schoknecht, & Carter, 1987; Forster & Davis, 1991). For the most part, priming experiments on form similarity

have not directly addressed neighborhood structure but have concentrated on the related though simpler question of whether formally similar words (e.g., *cat-can*) inhibit one another in perception.

As in the category retrieval literature, similarity effects in visual word recognition were, until recently, viewed as facilitatory (e.g., Hillinger, 1980; Meyer, Schvaneveldt, & Ruddy, 1974; see O'Seaghdha et al., 1992, for a review). However, Colombo (1986) found that high frequency words were inhibited by similar primes, whereas low frequency words were facilitated, and this phenomenon is now well established (see Lupker & Williams, 1987; Peterson et al., 1989).

Although interactive activation models have the potential for modeling frequency-dependent, similarity-based inhibition, models such as McClelland and Rumelhart (1981) do not. Rather, they postulate across-the-board, within-level lateral inhibition, which serves the function of segregating relevant words from competitors. Modifications of interactive activation models that could account for frequency-dependent inhibition between similar words can, however, be easily envisaged. Colombo (1986) proposed that lateral inhibition only applies between words that exceed a certain level of activation. However, this suggestion is rather ad hoc and undercuts the role of lateral inhibition in promoting perceptual sharpening (O'Seaghdha et al., 1992). Our earlier formal analysis suggests a simpler alternative. If lateral inhibition among similar words is learned, then the relation of inhibition to word frequency is a natural result. A high frequency word is more likely to be activated when it should not be, promoting the growth of inhibition. This analysis is compatible with Segui and Grainger's (1990) discussion of relative prime and target frequency effects and also comports with the evidence that low frequency but not high frequency words show net facilitation in high density neighborhoods (Andrews, 1989). The benefits of having congenial neighbors are balanced by the need to build fences in the case of high frequency words.

However plausible this scenario may be, the evidence suggests that acquired inhibitory connections may not in fact account for the empirically observed inhibition of high frequency words. Peterson et al. (1989) showed that when primes are forward masked, high frequency target words are facilitated rather than inhibited. If lateral inhibition between *cat* and *can* were acquired through experience, it should not be affected by a display manipulation such as masking. Peterson et al. also found that nonhomographic homophones (e.g., *muskel-muscle*) produce facilitation rather than inhibition. Learned similarity-based inhibition would dictate that homophones should accrue more lateral inhibition than other words, but in fact they produce facilitation rather than inhibition. Both results suggest that something other than lateral lexical inhibition—for example, competition for selection at the phonological

level (Peterson et al., 1989)—determines when behavioral inhibition is observed. If so, the differential receptivity of high and low frequency words to their neighbors (Andrews, 1989) may have a simpler explanation, such as that high frequency words are quickly recognized and therefore do not benefit from orthographic feedback.

Before leaving this topic, we briefly consider one other case, the inhibitory effect of phonetic similarity in auditory word perception recently discovered by Goldinger et al. (Goldinger, Luce, & Pisoni, 1989; Goldinger, Luce, Pisoni, & Marcario, 1992). In these studies, phonetically similar words are closest neighbors (defined in terms of a global index of the confusability of their segments); however, they do not share any phonological segments (e.g., *bull-veer*). Goldinger et al. contrast this kind of similarity with the kind of componential phonological similarity previously discussed (e.g., *bull-beer*). In auditory word identification and lexical decision, they find what appears to be a genuine inhibition effect for the phonetically related words but not for the phonologically related words. This is an interesting result, but we do not consider it to be compelling evidence for lateral inhibition. First, like other inhibitory effects, it could be accounted for by decision processes rather than by lateral inhibition (see Goldinger et al., 1992, p. 1233). That is, interactive activation models without lateral inhibition could easily account for the effect. Second, the effect appears to be specific to the auditory modality, and it is evident only when the words are presented in noise. This suggests that the effect may be due to the confusability of the phonological constituents of primes and targets rather than to the proximity of the words in an amodal lexical neighborhood.

Thus far, we have found good behavioral evidence for task-dictated competition based on similarity. This suggests that the conditions for the growth of lateral inhibition are present. However, we have found little unqualified support for the existence of lateral inhibition, itself, in language processing. In production, lateral inhibition may occur in the mapping from concepts to words, but effects of lateral inhibition can also be accounted for as competition for selection. The case for lateral inhibition among formally similar words in perception, as we have just discussed, is also weak. Our conclusions do not, of course, rule out the existence of lateral inhibition, but we suggest that its role in explaining similarity-based behavioral inhibition is limited.

In general, resolution of the status of intralevel lateral inhibition depends on the ingenuity of experimenters in defining experimental tests that both address identifiable levels and test directly for inhibitory mechanisms. For example, Frauenfelder, Segui, and Dijkstra (1990) conducted a direct test of lateral inhibition at the phonemic level of Elman and McClelland's (1986) TRACE model of speech perception. Their conclusions were negative. Likewise, Marslen-Wilson (in press) reports

a test of lexical-level inhibition in auditory word recognition as a function of the density of lexical neighborhoods. Marslen-Wilson argues that the evidence speaks against lateral inhibition, but in favor of bottom-up inhibition. We cannot evaluate these studies in detail here. Rather, we cite them as examples of the kind of work that may eventually resolve the status of lateral inhibition.

3. Reverse Effects In the absence of strong evidence of task-dependent lateral inhibition in production or perception, the status of the reverse effects we considered earlier is even more difficult to evaluate. The reverse effects are semantic inhibition in perception and form-related inhibition in production. We now consider each of these in turn.

To our knowledge, the work of Dagenbach and Carr (Dagenbach, Carr, & Wilhelmsen, 1989; Carr & Dagenbach, 1990; Dagenbach & Carr, this volume) provides the only well-documented case of semantic inhibition in perception. Briefly stated, Dagenbach et al. made the nonintuitive discovery that the nature of the task used during a threshold-setting procedure influenced the effect of masked semantically related primes in subsequent priming trials. Specifically, when semantic similarity judgments were made during the threshold setting procedure, semantically related masked primes subsequently produced inhibition (see Dagenbach & Carr, this volume). This is not the kind of reverse effect our theoretical analysis requires however. According to Dagenbach et al. (1989), it is more akin to the kind of retrieval inhibition we discussed in the context of production than to the kind of hardwired lexical inhibition that might arise in a shared architecture for perception and production. The threshold-setting procedure appears to influence the operation of early retrieval processes, the same retrieval processes that are involved in mapping from concepts to words in language production, and the inhibitory effect appears to take place at this level rather than at the level of word representations. In addition, the effect, like other inhibitory effects on retrieval we have considered, is dynamic and transient rather than hardwired.

The other reverse case we are interested in is form-related inhibition in production. Given our negative conclusions with regard to form-related lateral inhibition in perception, it would be surprising to find clear evidence in production. However, there are several relevant cases to consider.

Meyer and Gordon (1985) found that similar phonemes inhibit in production, a potential instance of reverse similarity-based inhibition. They attributed the effect to lateral inhibition between similar phonemes that arises from bottom-up feedback from features to phonemes. When planning to produce a particular segment, say /b/, one needs to inhibit similar segments such as /d/ that are activated by feedback from

shared features. Thus, they suggest that the effect is due to feedback at the phonological level. In this sense, Meyer and Gordon are opting for true lateral inhibition and, particularly, for an explanation of reverse effects as in our half-modular architecture, where the relevant levels are the phonological feature and the phonological segment, rather than segments and words. This is an appealing account and is, as well, the only instance in the literature of an explanation of molar-level reverse inhibition in terms of lateral inhibition needed to overcome bottom-up feedback. However, as with other simple molar effects of inhibition, there is no compelling need to postulate acquired inhibitory connections.

Sevold and Dell (submitted) have shown that the production of phonologically similar word pairs has an inhibitory component if the similar sounds are word-initial. However, they attribute the inhibition to competition at the phonological level (in the spirit of Peterson et al's, 1989, explanation of form-related priming), rather than to lateral inhibition between formally similar word nodes. There are several other examples of form-related inhibition in production. For example, it has been suggested that tip-of-the-tongue states may be induced by form-related competitors (Jones & Langford, 1987; Woodworth, 1938), but Meyer and Bock (1992) and Perfect and Hanley (1992) have shown that this conclusion may be premature. Bock (1987), Levelt et al. (1991), and Dell and O'Seaghdha (1992) discuss several more complex cases. However, the examples we have considered are sufficient to demonstrate that there is little basis for attributing behavioral form-based inhibition in production to lateral inhibition.

Overall, our review of paradigmatic inhibition reduces to two key points. An abstract analysis shows that, in principle, similarity-based lateral inhibition could be acquired from experience. A survey of experimental studies, however, finds no clear-cut demonstration of the existence of such inhibition. The cases of molar or behavioral inhibition we have reviewed either have an alternative explanation or can be accounted for by interactive activation models that do not involve lateral inhibition. Lateral inhibition is still a viable candidate explanation in the latter cases. Demonstration of its existence, however, remains a challenge for experimenters.

V. SYNTAGMATIC PROCESSES

Syntagmatic decisions involve the serial order of linguistic items. In activation-based models of production, these decisions require the activation in correct sequence of the units representing the items. First, we consider noncreative syntagmatic decisions in which the sequence is already known; for example, the order of the sounds in a word or the

order of words within a formulaic or idiomatic phrase. Then, we turn to the more complex issues of creative sequencing.

A. Noncreative Syntagmatic Decisions

The earliest accounts of sequencing, *chain associative* models, used forward excitatory associations. Each item in a sequence excited its successor. For example, in producing the word *cat*, /k/ excites /æ/, which then excites /t/. The problems with the simplest kind of chain association are well known (Lashley, 1951; see MacKay, 1987, for review). If all of the sequences make use of the same elementary units, the associative chains are completely lost: /k/ not only has an association to /æ/, but to lots of other sounds as well. As a result, there is no way to choose which association to follow. Hence, there has to be some kind of plan or goal that controls which associations are followed, or which units are considered. We consider two proposals for how such plans might be represented, Estes' (1972) hierarchical scheme and Jordan's (1986) recurrent network.

1. Estes' Hierarchical Model Estes (1972) based his theory of the long-term representation of serial order on a hierarchical organization. Contiguous elementary response units form chunks, and contiguous chunks form higher level chunks (see Figure 3). A chunk unit such as *cat* tends to excite its constituents, /k/, /æ/, and /t/. Each constituent unit inhibits all of the others that follow it, for example, /k/ inhibits /æ/ and /t/.

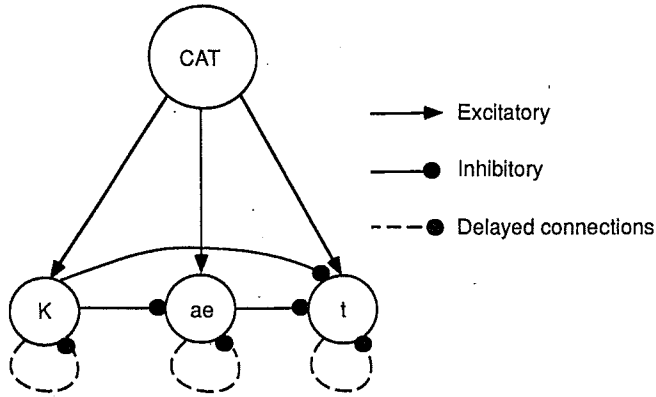


FIGURE 3
The Estes model of the long-term storage of serial order. CAT is the chunk or plan unit. Its activation triggers the sequential production of its three sounds.

Moreover, each unit enters a period of self-inhibition after it has become sufficiently active to be produced.

Let us consider how serial order is regulated by this scheme. First, *cat* begins to activate /k/, /æ/, and /t/; /t/, and to a lesser extent /æ/, are inhibited by earlier units and, hence, initially only /k/ sustains its activity. After /k/ reaches a predetermined threshold and is produced or promoted to the next level of processing, it undergoes self-inhibition. As a result /æ/ is disinhibited, becomes the most active constituent, and is in turn selected. Finally, /t/ is selected in the same manner.

The Estes scheme is admirably simple and can explain a number of findings in serial learning. In particular, it offers an account of why chunk sizes of three lead to better retention than others. Organizing a sequence into chunks of three requires fewer connections within and between the chunks.

2. Jordan's Recurrent Network Another general solution to representing stored syntagmatic relations was offered by Jordan (1986). The scheme includes both a plan and a context that changes throughout the execution of the plan, and allows the plan and the context to combine in a nonlinear fashion in determining what unit is to be produced next. Figure 4 shows the design of the system, which contains three layers: an input layer, which includes the plan and the context, a hidden layer, and an output layer. As in many models of this sort, the activation value of each node is a sigmoidal function of its inputs. The figure shows an illustrative network with three possible plans, the words *cat*, *act*, and *tack*, each associated with a single input unit. There are three output units, one for each segment. Connections can be either excitatory or inhibitory and run from the plan units, through the hidden units, to the output units. There are no within-level connections and no interactive feedback connections in the conventional sense, so the architecture is at least superficially very different from an interactive activation network. The connection weights are set by a learning algorithm such as back-propagation. The novel component of Jordan's approach is the addition of a set of context or state units that serve, technically, as input, but are actually derived from the previous output of the network (see Figure 4). Each context unit's activation is a function of that of a corresponding output unit and its own previous level of activation.

Serial order is produced because the pattern of activation among the context units changes as the sequence is produced, and this pattern can combine nonlinearly with the plan to control the output. When the plan is to produce *cat*, the plan units are activated appropriately, say 1 unit of activation for *cat*, and none for the other plan units, and the context units have no activation. Under these circumstances, the model's connection weights are tuned by the learning process, so that activation

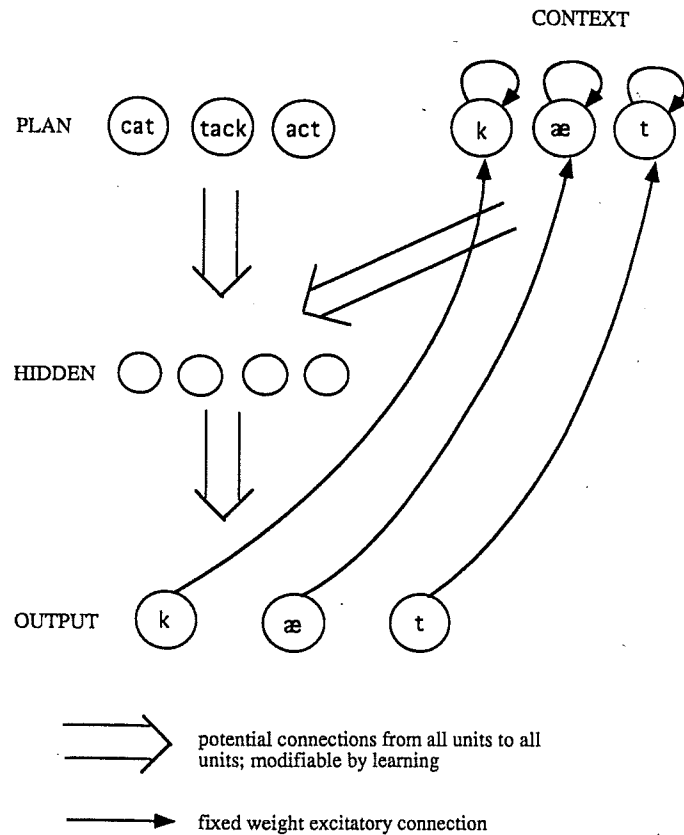


FIGURE 4

A Jordan network for producing *cat*, *tack*, or *act*.

flows from the input to the /k/ output unit. Having produced the first sound, each output unit sends activation to its corresponding context unit. Assuming that the output was in fact /k/, the context unit for /k/ acquires some activation. At this point, there is another forward spread of activation from input to output. Although the plan has not changed (the *cat* plan unit is still active), the context is different from what it was at the beginning of the word. The network “knows” that it has just produced /k/, and therefore it now learns to produce /æ/. Thus, as long as the context keeps changing, the network knows where it is in the sequence.⁴

⁴Even if there are repeated items in the sequence, no problem arises because the context is not just a copy of the previous output, it is a function of the previous output and the previous context.

The Estes and Jordan systems are similar in that both embody mechanisms for relating a stored plan to a sequence by means of spreading activation through excitatory and inhibitory connections. Also, both embody a hierarchical organization. An element produced in a sequence can, itself, be a plan for another constituent of the sequence—an essential feature for a language production model. In the Estes model, there are two particular kinds of inhibition: asymmetric lateral inhibition, in which each output unit inhibits those following it, and delayed self-inhibition, the tendency for each node to turn itself off after it has been fully activated. These two functions are accomplished in Jordan's model by the way the activation on the context units combines with that of plan units. But the actual mechanisms—for example, the extent to which inhibition is used—are not stipulated in the model. The learning process determines the configuration of excitation and inhibition. Therefore, the extent to which inhibition is used depends heavily on the structure of the set of sequences that is learned.

3. The Generalized Estes Model Let us consider the relation between the two models in more detail. Figure 5 shows a three-unit network that illustrates what we call the generalized Estes model. It consists of a plan unit for the sequence AB, a response unit A, and a response unit B. The network has top-down connections from the plan unit to each response unit, and has two kinds of lateral and self-connections—immediate and

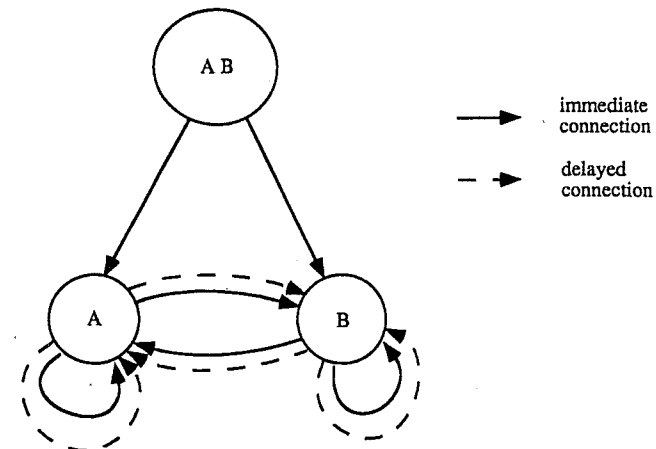


FIGURE 5

The generalized Estes model. Activation of the plan AB results in the activation of A and the inhibition of B, then the activation of B and inhibition of A, and finally the inhibition of both A and B. Excitation versus inhibition is not specified.

TABLE 1
List of Connections in the Generalized Estes Model

AB→A	Chunk to first unit
AB→B	Chunk to second unit
A→A	Immediate self-connection for first unit
B→B	Immediate self-connection for second unit
A→B	Immediate forward lateral
B→A	Immediate backward lateral
A→A(d)	Delayed self-connection for first unit
B→B(d)	Delayed self-connection for second unit
A→B(d)	Delayed forward lateral
B→A(d)	Delayed backward lateral

delayed. Immediate connections send activation at time step t , based on the activation of the source node at $t - 1$. Delayed connections send their activation based on the source node's activation at some earlier time step. The entire scheme has 10 potential connections, which are listed and labeled in Table 1.

The serial order of a two-unit chunk can be produced in several ways using this generalized model. For example, a simple forward chain association model might have excitatory connections AB→A and A→B(d), and delayed self-inhibition A→A(d) and B→B(d). Assuming that the delayed connections have a lag of four time steps, and that one starts by activating AB, this configuration will activate A and then after four time steps, turn A off and activate B, which will, itself, turn off after four time steps. At this point, the plan has been executed and AB is turned off.

The original Estes approach can be characterized by the following constraints on the connections: AB→A and AB→B are excitatory and equal in strength (symmetric chunk connections); A→B is inhibitory (immediate forward inhibitory connection); and A→A(d) and B→B(d) are inhibitory (delayed self-inhibition). It is also in the spirit of Estes' discussion that activation tends to persist in a node over time. This can be represented by immediate self-excitation (A→A and B→B are excitatory). All other connections are null. We can make this more concrete by working through an example. Assume that at each time step, each node i possesses an activation value $A(i)$ between 0 and 1 with a resting level of .5, and that the activation of a node is a sigmoidal function of its input. Specifically, for each time step,

$$A(i) = 1/(1 + \exp(-\text{input } i))$$

and

$$\text{input } (i) = \sum_j A(j)w_{ij}$$

where w_{ij} is the connection strength from node j to node i .

TABLE 2
Production of a Two-Item Sequence by the Generalized Estes Model

Time step	Activation of A	Activation of B
1	.88	.88
2	.88	.18
3	.94	.10
4	.94	.07
5	.18	.03
6	.09	.72
7	.04	.90
8	.03	.93
9	.55	.94
10	.30	.07
11	.51	.09
12	.57	.03
13	.10	.02
14	.20	.37

Connection weights: AB→A, AB→B = 2
 A→B = -5
 A→A, B→B = 1
 A→A(d) = -5
 B→B(d) = -1

Suppose that the goal is to turn on A and then B, where "turn on" means that a node has an activation of greater than .6 for four time steps in a row. If we assume that the chunk node is given an activation level of 1 for nine time steps and that the delay connections send activation after a lag of four time steps, the set of connection weights shown in Table 2 will achieve the goal: A is activated for steps 1 to 4, and B is activated for steps 6 to 9. After that, nothing turns on, but the activation levels continue to oscillate with diminishing amplitude.

The generalized Estes model can be shown to be a kind of Jordan network by placing the nodes in the Jordan configuration shown in Figure 6. The AB node is a plan unit, A and B are output units, and a, b, a(d), and b(d) are context units. As in the basic Jordan network, the context units derive their activation from the output and they then serve as input during the next time step. The a and b context units copy the activations of the A and B units, respectively, and the a(d) and b(d) context units store the activations of the A and B units four time steps earlier. A further characteristic of this network is that there are no hidden units—each input unit (the plan and context units) has a potential direct connection to each output unit. Thus, the network's 10 input-output connections correspond to the 10 connections presented in Table

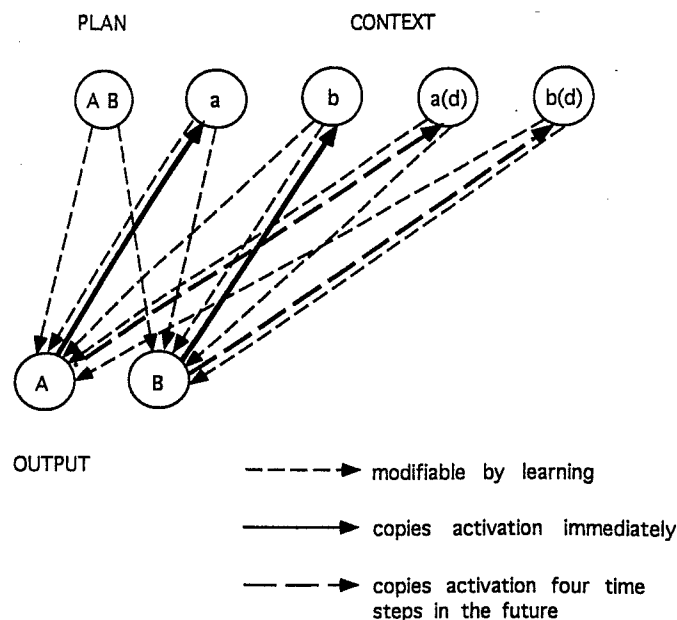


FIGURE 6

The generalized Estes model expressed as a Jordan network with both immediate and delayed context units. Excitation versus inhibition is not specified.

1 and Figure 5. The connections from the *a* and *b* context units represent either immediate lateral connections (as in $a \rightarrow B$ or $b \rightarrow A$) or immediate self-connections ($a \rightarrow A$ or $b \rightarrow B$). Those from the *a(d)* and *b(d)* context units to the output units are the analogous delayed connections. The two configurations are computationally equivalent when it is further stipulated that a single time step involves the spread of activation from input to output and the updating of the context units' activations.

4. Learning in the Generalized Estes Model We are now in a position to use the learning aspect of the Jordan network to investigate how the inhibitory configuration of the Estes approach might be acquired. In the section on paradigmatic decisions, we speculated that lateral inhibition might grow in response to an error-correcting learning rule. Jordan's network sets connection weights by this kind of rule. We will use the simplest version of an error-correcting rule: The change in weight from an input unit *i* to an output unit *j* is

$$\epsilon(t_j - a_j)a_i$$

where ϵ is the learning rate, t_j is the target activation of an output unit,

a_i is its actual activation, and a_i is the input unit's activation. We set ϵ to 0.1, a moderately small value. Small values are associated with a greater probability of eventual success in learning but have, of course, the disadvantage of requiring many trials.

The original Estes model used immediate forward lateral inhibition and delayed self-inhibition. To find out if these inhibitory relations would be learned, we started by giving the network in Figure 6 the task of learning to produce *A* for four time steps (target activation of *A* = 1 and *B* = 0), *B* for four time steps (*A* = 0 and *B* = 1), and then neither for four more steps (*A* = 0 and *B* = 0). The *AB* unit was turned on for the first eight time steps. The network was required to fit the original Estes mold in that its $AB \rightarrow A$ and $AB \rightarrow B$ connections were fixed at a weight of +2.0 and it could have no lateral delayed connections [$a(d) \rightarrow B$ and $b(d) \rightarrow A$ were fixed at 0]. Further, each output node was given a small persistence of activation by means of fixed immediate self-excitation; $a \rightarrow A$ and $b \rightarrow B$ were fixed at 1.0. Only the immediate lateral connections and delayed self-connections were modifiable by the learning process. After 200 trials, the network learned to produce the sequence about as well as the generalized Estes network configuration shown in Table 2, and it did, indeed, use asymmetric lateral inhibition ($a \rightarrow B = -4.25$ and $b \rightarrow A = -2.55$), and delayed self-inhibition [$a(d) \rightarrow A = -7.30$ and $b(d) \rightarrow B = -.83$].

Would the model discover the original Estes configuration if it were not constrained to have identical top-down connections and no delayed lateral connections? The answer to this question is negative. If all connection weights are initialized to zero and are modifiable by learning, the network becomes extremely accurate, much more so than in the constrained case, and its weight configuration exhibits the following variety of serial order devices (see Table 3):

1. Asymmetric top-down connections: The chunk node favors the first item by exciting it and inhibiting the second item.
2. Immediate lateral inhibition: Each response node tends to turn the other off.
3. Excitatory chaining: The first response tends to turn on the second one at a delay.
4. Backward delayed inhibition: The second response turns the first one off at a delay.
5. Delayed self-inhibition: This device is the same as in the original Estes model.

One other important feature of the unconstrained model is that the weight pattern shown in Table 3 is stable. That is, regardless of the initial values of the weights, the learning process arrives at a similar configuration. For example, if the weights are initialized to the same

TABLE 3
Results of Learning to Produce a Two-Unit Sequence

Time step	Activation of A	Activation of B
1	.90	.29
2	.94	.06
3	.96	.07
4	.96	.07
5	.13	.82
6	.03	.95
7	.02	.95
8	.02	.96
9	.01	.06
10	.02	.07
11	.02	.07
12	.02	.07

Connection weights (after 100 trials):

Plan to output:	$AB \rightarrow A = 2.17$; $AB \rightarrow B = -0.93$;
Immediate self-connections:	$a \rightarrow A = 1.10$; $b \rightarrow B = -1.19$
Immediate lateral connections:	$a \rightarrow B = -1.66$; $b \rightarrow A = -1.76$;
Delayed self-connections:	$a(d) \rightarrow A = -4.42$; $b(d) \rightarrow B = -2.78$
Delayed lateral connections:	$a(d) \rightarrow B = 5.58$; $b(d) \rightarrow A = -3.84$

magnitudes but are assigned to the opposite directions to those shown in Table 3, learning still brings them back close to the configuration shown in the table.

5. Different Orders of the Same Elements We have shown that the Estes model can be expressed as a special kind of constrained Jordan network. In addition, some of the features of an Estes model, particularly its use of inhibition, can be acquired by the kind of learning that characterizes a Jordan network. We now turn to the real challenge for serial order models: Can these networks handle the serial order relations required in a model of language production? It turns out that neither the original nor the generalized Estes model is capable of storing the non-creative syntagmatic relations required to produce the phonological segments of words.

If we assume that phonological segments are response units, there are many cases where anagramlike permutations of the same elements comprise several different words. As MacKay (1987) pointed out, Estes' original model fails in these cases. For example, in the case of the words *oat* and *toe*, the order of /t/ and /o/ is indeterminate. The problem can be understood by recalling that the plan unit in the Estes model only activates a set of constituents, leaving the constituents, themselves, to sort out their order. Consequently, if there is more than one order associated

with the set, the model fails. When /o/ and /t/ are activated, each desperately tries to inhibit the other: /t/ attempts to inhibit /o/ because of the existence of *toe*, and /o/ does the same to /t/ because of *oat*. One might think that the problem could be solved by relaxing the constraint that the plan signal each constituent equally (identically weighted top-down connections), as we did with the generalized Estes model. So, for example, perhaps the connection from *toe* to /t/ would be stronger than that from *toe* to /o/, and the reverse would be true for *oat*. Unfortunately, this will not work either. If we attempt to learn a set of weights in the generalized framework from two plans, AB and BA, to the sequences, A B stop, and B A stop, respectively, the weight changes never make progress toward a solution. It is not difficult, though, to learn weights to store other pairs of sequences, such as AA and BB, or AB and BB. It is the anagram case that presents the problem.

The difficulty that both the original and the generalized Estes models experience with storing different sequences of the same elements is related to the well known intractability of the exclusive OR (XOR) problem in a two-layered network (Minsky & Papert, 1969; Rumelhart et al., 1986). The XOR problem is the classic example of a mapping that requires the computation of an interaction or nonlinear combination of inputs. The expression $p \text{ XOR } q$ is false when p is false and q is false, and when p is true and q is true. If one identifies "true" with an activation of 1 and "false" with 0, and sets up a network of two inputs, one each for the truth values of p and q , and one output unit that is supposed to compute $p \text{ XOR } q$, one cannot find a set of weights that performs the computation (assuming that output activation is monotonically related to net input). This can be illustrated geometrically by imagining an input plane whose dimensions are the activation of the input units p and q . Two points in the plane (0,0) and (1,1) require an output of 0, and two points (0,1) and (1,0) require an output of 1. The intractability of the problem is seen in the fact that the points associated with an output of 0 are not in a region of the plane that can be separated by a straight line from the region associated with an output of 1. When inputs requiring different output values of the same output unit can be separated by a straight line (or plane or hyperplane in higher dimensional input spaces) and this is true of all outputs, the problem is *linearly separable*. Problems such as XOR are therefore not linearly separable. In general, two-layered networks can only store sets of input-output pairs in which similar inputs lead to similar outputs. The input patterns that should turn on a given output unit must fall in a region of input space that is linearly separable from the ones that should turn that unit off.

The difficulty of the generalized Estes model with anagrams arises because it is a two-layered model—it has no hidden units—and the task of dealing with both AB and BA is not, as we have defined it, linearly

TABLE 4
The Problem with Anagrams^a

	Input units			
	AB	BA	a(d)	b(d)
Output unit for B should be ON				
Beginning of BA	0	1	0	0
Middle of AB	1	0	1	0
Output unit for B should be OFF				
Beginning of AB	1	0	0	0
Middle of BA	0	1	0	1
End of AB	0	0	0	1
End of BA	0	0	1	0

	Solution with hidden unit					Net input to B
	AB	BA	a(d)	b(d)	ABa(d)	
Connection weights to B	-	+	-	--	+++	
Output for B should be ON						
Beginning of BA	0	1	0	0	0	+
Middle of AB	1	0	1	0	1	+
Output for B should be OFF						
Beginning of AB	1	0	0	0	0	-
Middle of BA	0	1	0	1	0	-
End of AB	0	0	0	1	0	--
End of BA	0	0	1	0	0	-

^aBeginning = prior to time step 1; middle = prior to step 5; end = prior to step 9.

separable. Table 4 provides some insight into the task. We can specify the input space in terms of four dimensions corresponding to the two plan units, AB and BA, and the two context units a(d) and b(d).⁵ We focus only on the output unit B and consider when it should be on. The table presents six moments defined by three locations in the input patterns AB and BA. To illustrate the difficulty, let us attempt to determine which connections from the input units to B should be excitatory and which should be inhibitory. First, the connection BA→B must be excitatory because BA is the only input unit on at the start of a BA sequence. Similarly, AB→B must be inhibitory to turn B off at the beginning of AB. Finally, both a(d)→B and b(d)→B must be inhibitory to turn B off at the

⁵For simplicity, we are not considering the a and b immediate context units. Only certain key times in the production of the sequences are being examined and, at these times, the activation of a equals a(d), and b equals b(d). Consequently, the immediate context units are redundant and do not expand the input space.

end of BA and AB sequences, respectively. However, these connections don't work for the middle of AB, where the net input from the connections that we have specified thus far is strongly inhibitory and it should be excitatory, and for the middle of BA where the net input is neutral and it should be inhibitory. The latter difficulty can be fixed by making b(d)→B doubly inhibitory so that the excitatory contribution of BA→B is overwhelmed. But there is no such quick fix for the middle of AB. In essence, this is the multidimensional analogue of the XOR problem.

The problem can be solved by adding another unit, one that only turns on when AB and a(d) are on. Then, a triply strong excitatory weight from the new unit to B will permit B to correctly turn on in the middle of an AB sequence and, generally, will allow B to be turned off and on when it should. This new unit thus must obtain activation from the input units and, in turn, send activation to the output unit B. So, it is effectively a *hidden* unit. Adding one or more hidden units to a two-layered model is the only way to implement a mapping that is not linearly separable.

The standard Jordan network, unlike the generalized Estes model, is set up with a layer of hidden units just so that complex sets of sequences can be stored. Although the actual use of hidden units will be determined by the patterns to be learned and architectural constraints, the hidden units will come to encode nonlinear combinations of inputs. It is as if the plan unit in the Estes model not only signals which items are in the sequence, but also tells which particular connections among the items should be followed for a particular plan.

One very interesting way that an extra layer of units can be used is as a competitive filter. Houghton's (1990) competitive cueing model uses three layers of units: plan and response units analogous to what we have discussed thus far, and a third layer that uses very strong lateral inhibition to resolve competition. Specifically, suppose that there are two plan units, AB and BA, connecting to two response units, A and B. In Houghton's model, the plan units use asymmetrical top-down excitatory connections to simultaneously retrieve all of the responses for a given plan, such that response activation is a function of position in the plan. For example, the first response is more active than the second. The activation from the response units is then passed on to the competitive filter layer, which, for our simple example, would contain units A' and B'. A' would receive excitatory input from A, and B' from B. In the filter layer, A' and B' strongly inhibit each other so that only one of these units, the one whose response unit had the most activation, is activated and the other is inhibited. This constitutes the selection of the first unit. Then, the single activated filter node strongly inhibits its corresponding response unit with the effect that it is no longer activated. At this point, the most highly activated response unit is the one that was next most

strongly signalled by the top-down connections and so the competitive filter will then select this next unit, and so on.

Notice how the model solves the anagram problem. AB would initially excite both A and B, but A would be more activated. There is no need for either A or AB to inhibit B because this is accomplished in the filter layer, where A' ends up with all the activation and B' is inhibited. Then, after A' inhibits A, B' will win in the filter, leading to the production of B. Finally, B is inhibited by B' shutting everything down. If BA is the plan, everything works the same way except that the top-down connections favor B over A. In general, the competitive cueing model works because it separates the competitive interactions (in the filter) from the retrieval of the items of the plan (in the response units). This makes it possible for order to be stored in the connection weights from plan to responses. Although the competitive cueing model is very different from Jordan's model (e.g., it uses different learning rules), there is, we believe, also a deep similarity. Both allow for the storage of any order because the decision to output an item at a particular time is determined by a nonlinear combination of plan information (plan-to-response connections in Houghton's model) and contextual information (context units in Jordan's model, the activated state of the response and filter units in Houghton's model).

Variations on Jordan's network, particularly those developed by Elman (1989, 1990), have been used with some success in accounting for the facts associated with the storage of phonological sequences, facts about coarticulation (Jordan, 1986), phonological speech errors (Dell et al., 1993), and other linguistic patterns (e.g., Corina, 1991; Gasser & Lee, 1990; Hare, 1990). As we turn to creative syntagmatic relations, however, we will see that the applicability of these kinds of networks to language is controversial.

B. Creative Syntagmatic Relations

A central feature of any language production model is that it must allow for the production of novel sequences. The creative component to language is most obvious in the need to combine words to produce novel sentences, although the ability to build new words out of existing morphemes and phonological segments is an important part of the creativity in many languages.

Activation-based systems that retrieve order from chunks, such as the generalized Estes model, are limited to reproducing learned orders and, hence, are not equipped to create novel orders. As we have just discussed, the problem is that these systems attempt to order content directly by stipulating connections in advance. For example, the chunk

big-dog must be stored in memory if *big* and *dog* are to be produced in sequence. To overcome this restriction, the schemata, rules, or structures that regulate order must be represented separately from the content items that they operate on. So, rather than directly storing the sequence *big-dog*, there needs to be an independently represented rule that says that things in the category ADJECTIVE come before things in the category NOUN. Then, it is possible to produce both familiar (*big dog*) and unfamiliar (*tiny aardvark*) content sequences. The only serial order that need be stored is the order of the linguistic categories.

The separation of linguistic rules or structures, which bear the primary responsibility for order, from linguistic content is a major tenet of linguistic and psycholinguistic theory. Every model of language production, including those that are based on spreading activation, employs structural frames or other generalizations that operate on linguistic categories, not individual items (see Levelt, 1989, for a review). These frames enable the system to order items creatively. To illustrate, we now consider two related activation-based models. The first is MacKay's (1982, 1987) model and the other is a model that has emerged from the recent work of Berg, Schade, and Eikmeyer (e.g., Berg & Schade, 1992; Eikmeyer & Schade, 1991). The latter model is similar in many respects to the earlier model of Stemmerger (1985).

1. MacKay's Model MacKay's use of structural and content nodes in production is illustrated in the top of Figure 7. Although we focus only on MacKay's treatment of creative syntagmatic processes in language production, it should be recognized that the entire theory deals with perception as well as production, learning as well as performance, and, moreover, applies to nonlanguage phenomena as well. The model is a network with nodes organized along two dimensions. First, there are levels of processing corresponding to conceptual (including the word and phrase level), phonological, and motor levels. The figure shows only the conceptual level. Second, there is a distinction between *content nodes*, which are organized hierarchically and represent particular phrases, words, syllables, syllabic constituents, segments, and so on, *structural nodes*, which represent linguistic categories and ordered relations among the categories, and *timing nodes*, which control the rate of activation of the structural nodes. The figure shows only content and structural nodes.

The production of a sequence such as *the girl* starts with the activation of the content node for the entire phrase. One can think of this phrasal content node as representing the complex concept of GIRL (singular, definite). Many such complex concepts will already be in memory. Others, such as *green cows* will be the product of inferential processes (e.g., what happens when you drop a load of green paint from

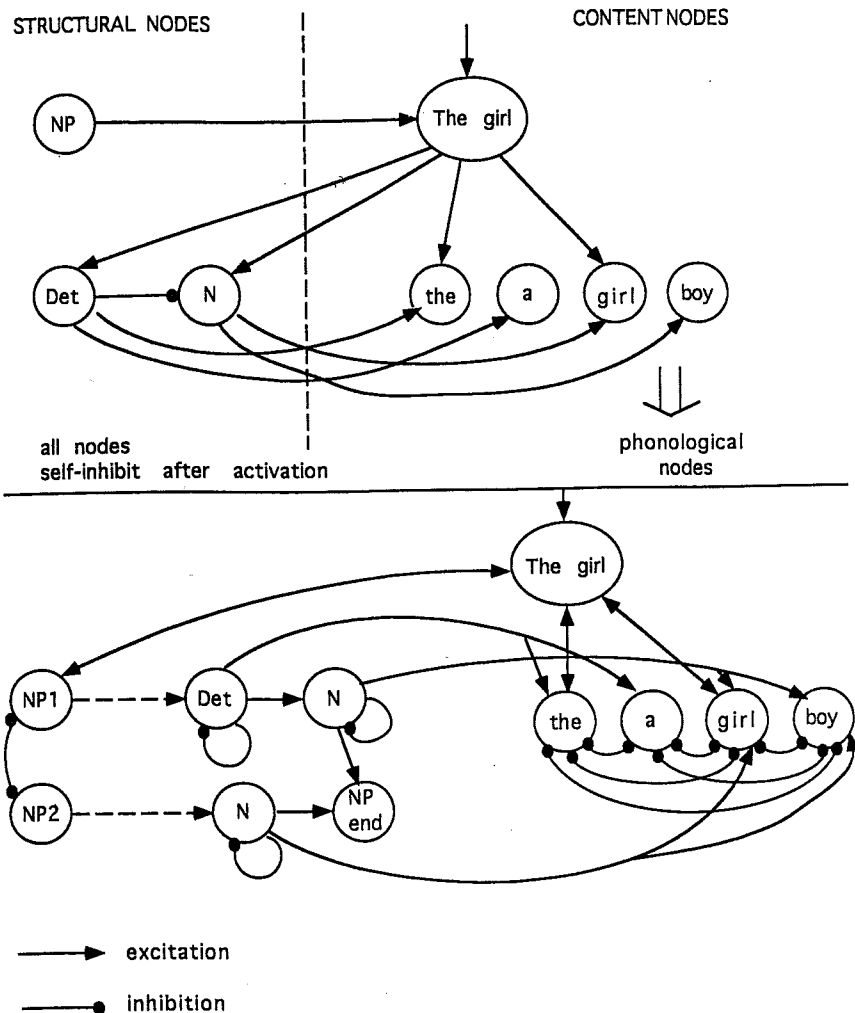


FIGURE 7
Models of creative syntagmatic decisions in production. MacKay (1982) top and Eikmeyer and Schade (1991) bottom.

a plane onto a dairy farm?) or perceptual processes (e.g., from the unlikely event of seeing green cows). Thus phrasal nodes are assumed to exist prior to the sequencing of their words. For our example, we will label the phrasal node as "the girl," even though it is not, itself, a sequence of words. When *the girl* is activated, it primes the content nodes for *the* and *girl*, and the structure nodes *det* and *N*. Recall that, as previously mentioned, MacKay distinguishes between the activation of a

node, which is characterized by a sustained and very high level of activation, and priming, which involves lower levels of activation. The structure nodes also receive priming from a timing node. Because *det* inhibits *N*, the net result is that priming builds up most quickly on the structure node *det* and it becomes activated. The activation of *det* primes the content nodes for all possible determiners and, in the case of *the*, the summation of priming from *det* and *the girl* is sufficient to activate it. After each node is activated, it enters a period of self-inhibition. So after *det* is activated, it is inhibited, and now the structure node for *N* can become activated when the timing node primes it. The activated *N* node then primes all the content nodes that it connects to, that is, all nouns, and *girl* becomes activated because of its extra source of priming from *the girl*.

There are several important characteristics of MacKay's approach from our perspective. One is that the relation between *the girl* and *det* and *N* in the model is just like that in the Estes model. Serial order is controlled by a hierarchical chunk and by immediate forward inhibition. Moreover, each node undergoes self-inhibition after it is activated. Does the model, then, experience the same difficulties that the Estes model does when there are different orders involving the same items? MacKay avoids this problem by means of additional features in his model. First, ordered relationships are not marked on the content nodes, but rather on the structural nodes. It turns out that there is much less need to worry about different orders of the same nodes at the structural level. For example, the only ordering involving determiners and nouns is *det*, *N*. However, this cannot be the entire solution because there are cases in which different orders of the same linguistic categories do occur. At the phonological level, for example, there are syllables composed of the categories C (consonant) followed by V (vowel), as well as the reverse. MacKay's solution is to treat the structural nodes not as strict linguistic categories such as vowel, consonant, adjective, or noun, but rather as sequentially defined categorical domains. For example, the category of C in the case of a CV syllable would belong to the domain onset, whereas the C in a VC syllable would belong to the domain coda.⁶ Hence, the same response units are not used. Some evidence that categories are sequentially defined comes from the fact that phonological speech errors tend to involve onsets replacing other onsets, and codas replacing codas (MacKay, 1970). So, although MacKay makes use of the Estes configuration, he does so in a way that avoids the difficulties.

The absence of lateral inhibition among competing alternatives at the content level is another feature of MacKay's model. As we mentioned

⁶MacKay's use of rhyme units intervening between syllables and segments also helps to avoid the problem; a rhyme unit dominates a coda, but not an onset.

earlier in our discussion of paradigmatic selection, this is compensated for by MacKay's notion of activation. When priming builds up in several nodes within the same categorical domain (e.g., *girl*, *boy*), it is stipulated that only the first one to reach an activation threshold becomes activated. And when a node achieves activated status, it is given an activation level much greater than what can be achieved by mere priming. To make this more concrete, assume that priming can build up to a level of 50 units of excitation. The first node in a domain to get to 50 becomes activated and immediately ups its excitation level to 100 units. Furthermore, every other node in the domain is then prevented from being activated and, hence, cannot exceed 50 units of excitation. The effect of this is very similar to what might be achieved by lateral inhibition within a domain. In general, MacKay's model, like models with lateral inhibition, has the property that few nodes can be highly active at once. Only one node from a domain can achieve an activated state and the sequential nature of the structure nodes largely ensures that nodes in only one domain at a time are building up appreciable levels of priming.

In sum, MacKay's model deals with creative syntagmatic decisions by a combination of noncreative syntagmatic mechanisms, such as the Estes sequencing schema, that operate on structurally defined categories, paradigmatic decisions among competing content items, and links between the two mechanisms. This solution is, in gross terms, quite similar to that adopted by other activation-based models. The model of Eikmeyer and Schade (1991) presented in the bottom of Figure 7 can be used to illustrate some dimensions of difference.

2. The Eikmeyer and Schade Model The first difference between Eikmeyer and Schade's (1991) model and that of MacKay (1982, 1987) is that the structural nodes are ordered by forward associative chains rather than by the Estes configuration. The content node for *the girl* excites the content nodes for *the* and *girl*, and it also excites an appropriate syntactic node for NP1. The syntactic node then excites the first category of the sequence, Det. Det sends activation to all possible determiners (as in the MacKay model) and to the next category N, and so on until a special NP-end node is reached. The sequencing of categories is achieved via a combination of forward excitatory connections and self-inhibition. The higher level structure nodes (e.g., noun phrases) do not stand in a hierarchical relation to their sequential constituents, but rather serve as starting states (see Dell, 1988, for a similar configuration). As we showed earlier, this kind of serial order device will not work when units participate in more than one chain. Eikmeyer and Schade, therefore, used different units for different chains, as shown in Figure 7 for det N and N.

Another difference between the two models concerns lateral inhibition. Eikmeyer and Schade use lateral inhibition between competing sequential chains, as shown in the figure by the lateral inhibition between NP1 and NP2, and between all content units at a particular level, as shown in the lateral inhibition among the word units. This inhibition ensures that only one sequence at a time controls behavior and that only one content item becomes highly activated. It gives lateral inhibition a role in both planning and producing a sentence. Not only do *girl* and *boy*, traditional paradigmatic competitors, inhibit one another, but also *girl* and *the* inhibit one another because they are words in the intended utterance. According to Berg and Schade (1992), this kind of arrangement allows for the planning or subthreshold activation of all of the words in an utterance. During planning, lateral inhibition keeps the activation of each intended word low, but during production inhibition is overcome when the syntactic nodes activate in sequence.

Thus, Eikmeyer and Schade's use of lateral inhibition serves the same functions that MacKay addresses by distinguishing between priming and activation, and by allowing only one node in a domain to become activated. Berg and Schade (1992) suggest that one can tell whether certain production decisions are truly implemented by inhibition or by other means, by attempting to model atypical production such as that of aphasic patients or children. The idea is that it may be possible to attribute certain phenomena to a general problem with inhibition, given a model that treats inhibition in a particular fashion. We agree that such data and modeling efforts would provide useful constraints on theory, but our earlier conclusion concerning the difficulty of identifying lateral inhibition in experimental data suggests that the attribution of a cluster of behavioral effects to an inhibitory problem will be quite difficult.

3. Evaluation of the Models Both MacKay and Eikmeyer and Schade use serial order devices for storing the order of sequences of categories that are less powerful than mechanisms capable of storing any orders (e.g., Jordan networks with hidden units), and both solve the problem by using different units for the same categories when they have different sequential functions. Which is better? The MacKay system has the advantage that it does not require a separate chain for every order. Moreover, it supports its use of categories by pointing to the speech error literature, in which sequentially defined domains account for what units replace one another. The forward chaining system of Eikmeyer and Schade, though, has the virtue of a clear implementation, one that does not require anything other than spreading activation.

We suggest that for the most part the two systems make similar empirical predictions. However, one may be able to distinguish between them by testing whether two related syntactic structures use common

nodes. The Eikmeyer and Schade system uses, for example, different chains for Det N and Det adj N, whereas in MacKay's system the same Det and N nodes could be used. The hierarchical nature of the Estes configuration would supply the Adj when it is part of the sequence. Syntactic persistence effects (e.g., Bock, 1986) could, at least in principle, reveal whether these related sequences use common sequential resources. For example, one could have subjects produce a particular priming syntactic structure and then see if there is any benefit or cost to later using other target structures. The idea is that phrases or sentences that are assumed to use the same structural nodes may exhibit a priming effect. By examining pairs of prime-target structures which one theory says use common nodes and the other theory says use distinct nodes, one could conceivably discriminate between the theories.

4. PDP Models To complete our discussion of creative syntagmatic relations in activation-based models, we briefly consider a quite different approach—an approach that questions the need for an explicit distinction between content and structural nodes. Researchers who adopt a parallel-distributed-processing (PDP) approach to connectionist modeling have developed models of language processing in which rules or other structural generalizations are not built into the architecture, but rather emerge from the superimpositional storage of the linguistic sequences that comprise the set of training items (e.g., Corina, 1991; Elman, 1989, 1990; Hare, 1990; Gasser & Lee, 1990; MacWhinney & Leinbach, 1991; Rumelhart & McClelland, 1986b; Seidenberg & McClelland, 1989; St. John & McClelland, 1990). In short, these models use networks that can store sequences, as a Jordan network does, but because of the use of distributed representations, the sequences are stored in such a way that the network has some ability to extend itself to novel sequences. Rumelhart and McClelland's (1986b) model of English past tense allomorphy is the classic example. The network learns a set of paired root and past tense forms of verbs. Each verb is associated with a distributed representation involving many nodes and connections that are shared with other verbs. Consequently, the weight changes associated with the learning of a particular verb are mingled with those associated with other verbs. The superimposition of the weight changes allows the network to act as if it has extracted general rules. When given a novel verb, the network can, to some extent, form the correct past tense. Yet there is no rule that is stored separately from the storage of the trained verbs. The network is able to create new output, that is, to generalize to novel forms, without an explicit structure-content distinction.

At present, it is too early to say whether the creative syntagmatic relations required for language production can be exhibited by this kind

of PDP model. The issues are quite complex and the work done in the area is still preliminary. One model of this type was specifically developed to account for psycholinguistic data in production (Dell et al., 1993). That model used a variant of a Jordan network to store the phonological forms of English words as sequences of phonological features. Because the network had hidden units it had no trouble storing the sequences, that is, it could effectively deal with the required noncreative syntagmatic relations. However, Dell et al. wished to see if the model could exhibit structural effects that in other models have been ascribed to structural nodes or frames. These effects involved two characteristics of phonological speech errors. First, speech errors can create novel phonological strings, but only novel strings that obey the phonotactic constraints of the language being spoken. Second, there are syllable structure effects whereby errors tend to involve sets of features that comprise syllabic constituents. After the model stored sets of English words, its performance was degraded by noise added to the connection weights to simulate interference from surrounding words and other processing losses. The noise resulted in the model producing "speech errors" and the errors did, indeed, exhibit the structural effects. Dell et al. concluded that at least some of the effects attributed to phonological frames or structural nodes may, instead, reflect the combined influence of the stored vocabulary. For our purposes, this raises the possibility that powerful noncreative syntagmatic mechanisms may, when combined with distributed representations, deal with creative effects. Of course, Dell et al.'s efforts apply only to phonology, the least creative aspect of language production. It remains to be seen whether syntax can be similarly treated.

VI. SUMMARY AND CONCLUSIONS

Inhibition serves many purposes in interactive activation models of language production. Some models use lateral inhibition to aid in paradigmatic selection, others use inhibition only to regulate sequencing (e.g., MacKay, 1987), and others use inhibition in the service of both selection and sequencing (e.g., Eikmeyer & Schade, 1991). Our survey found good evidence for molar-behavioral inhibition in paradigmatic selection. However, the mapping of these effects to molecular lateral inhibition was unclear. In most cases, alternative mechanisms could also account for the data. Therefore, it appears that experiments that are specifically designed to test for the presence of lateral inhibition will be required to constrain models and theory in this area.

Whereas the status of lateral inhibition in tasks such as lexical retrieval is unresolved, inhibitory mechanisms appear to be essential to

the complex sequencing requirements of language production. We explored the range of these requirements and of the inhibitory mechanisms that serve them in a variety of different models. Although the functional requirement for inhibition in syntagmatic processing is clear, there are no specific processes analogous to lateral inhibition that the models mimic. The models serve the purpose of clarifying functional requirements and thus supporting theory building. Thus, for example, we were able to identify the difficulty some models have with the requirements of being able to represent different sequences of the same items. Interactive activation models like MacKay's and Eikmeyer and Schade's address this problem by a variety of means including a separation of structure and content representations. These models employ a range of inhibitory and excitatory devices that should in principle be susceptible to empirical evaluation. Finally, the development of PDP models that dissolve the distinction between content and structure presents a new challenge to interactive activation models. The PDP approach suggests that at least some effects that have been ascribed to structure may instead be emergent from content. To the extent that this is the case, the division of structure and content in activation-based models must be modified, and this in turn will have implications for the role of inhibition in these models. The need for the variety of inhibitory processes we have discussed remains. However, it may be difficult to justify assigning particular inhibitory devices to content nodes, and others to a separate set of structural nodes.

ACKNOWLEDGMENTS

This work was supported by NSF-89-10546 and DC-00191 (NIH). The authors wish to thank Ulrich Schade, David Adams, Vic Ferreira, Anita Govindjee, Scott McGurrian, and Chris Sevald for helpful suggestions and Linda May for work on the manuscript.

REFERENCES

- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 804-814.
- Bates, E. A., & Elman, J. L. (1992). *Connectionism and the study of change*. CRL Technical Report 9202. Center for Research on Language, University of California, San Diego.
- Berg, T. (1988). *Die Abbildung des Sprachproduktionprozess in einem Aktivationsflussmodell*. Tübingen: Max Niemeyer.
- Berg, T., & Schade, U. (1992). The role of inhibition in a spreading activation model of language production: 1. The psycholinguistic perspective. Unpublished manuscript, University of Oldenburg, Germany.

- Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of category specific inhibition. *Memory & Cognition*, 11, 500-510.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355-387.
- Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language*, 26, 119-137.
- Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99, 150-171.
- Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, 29, 413-432.
- Bowles, N. L., & Poon, L. W. (1985). Effects of priming on word retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 272-283.
- Brown, A. S. (1979). Priming effects in semantic memory retrieval processes. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 65-77.
- Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 204-215.
- Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 341-350.
- Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 226-234.
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI*. New York: Academic Press.
- Corina, D. P. (1991). *Toward an understanding of the syllable: Evidence from linguistic, psychological, and connectionist investigations of syllable structure*. Doctoral dissertation, University of California, San Diego.
- Dagenbach, D., Carr, T. H., & Wilhelmsen, A. (1989). Task-induced strategies and near-threshold priming: Conscious effects on unconscious perception. *Journal of Memory and Language*, 28, 412-443.
- Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychological Review*, 93, 283-321.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124-142. Reprinted (1989). In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 136-165). Cambridge, MA: MIT Press.
- Dell, G. S., Julianio, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149-195.
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287-314.
- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611-629.
- Eikmeyer, H.-J., & Schade, U. (1991). Sequentialization in connectionist language production models. *Cognitive Systems*, 3(2), 128-138.
- Elman, J. L. (1989). Structured representations and connectionist models. *Proceedings of the 11th Annual Conference of the Cognitive Science Society*, Ann Arbor, MI.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 213-252.

- Elman, J. L., & McClelland, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161-190). Washington, DC: Winston.
- Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.
- Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language*. New York: McGraw-Hill.
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North Holland.
- Forster, K. I., & Davis, C. (1991). The density constraint in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, 30, 1-25.
- Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology*, 39A, 211-251.
- Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, 16, 77-91.
- Fromkin, V. A. (1971). The nonanomalous nature of anomalous utterances. *Language*, 47, 27-52.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133-177). San Diego: Academic Press.
- Gasser, M., & Lee, C.-D. (1990). Networks that learn about phonological feature persistence. *Connection Science*, 2, 265-278.
- Glushko, R. (1979). The organization and activation of orthographic information in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, 28, 501-518.
- Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 1211-1238.
- Grossberg, S. (1978). A theory of visual coding, memory, and development. In E. L. J. Leeuwenberg & H. F. M. Buffart (Eds.), *Formal theories of visual processing*. New York: Wiley.
- Hare, M. (1990). The role of similarity in Hungarian vowel harmony: A connectionist account. *Connection Science*, 2, 125-152.
- Harley, T. A. (1984). A critique of top-down independent levels of models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, 8, 191-219.
- Harley, T. A. (1990). Paragrammatisms: Syntactic disturbance or breakdown of control? *Cognition*, 34, 85-91.
- Hillinger, M. L. (1980). Priming effects with phonemically similar words: The encoding bias hypothesis reconsidered. *Memory & Cognition*, 8, 115-123.
- Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 287-319). London: Academic Press.
- Jones, G. V., & Langford, A. S. (1987). Phonological blocking in the tip of the tongue state. *Cognition*, 26, 115-122.

- Jordan, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp. 531-546). Hillsdale, NJ: Lawrence Erlbaum.
- Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language*, 32, 474-516.
- Kawamoto, A. H., & Zemblidge, J. H. (1992). Pronunciation of homographs. *Journal of Memory and Language*, 31, 349-374.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: Wiley.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The timecourse of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122-142.
- Loftus, E. F. (1973). Activation of semantic memory. *American Journal of Psychology*, 86, 331-337.
- Loftus, G. R., & Loftus, E. F. (1974). The influence of one memory retrieval on a subsequent memory retrieval. *Memory & Cognition*, 2, 467-471.
- Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122-147). Cambridge, MA: MIT Press.
- Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- Lupker, S. J., & Williams, B. A. (1987). *When do rhyming primes inhibit target processing?* Paper presented at the 28th Annual Meeting of the Psychonomic Society, Seattle, WA.
- MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, 8, 323-350.
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210-227.
- 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: Springer.
- MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40, 121-157.
- Marslen-Wilson, W. (in press). Issues of process and representation in lexical access. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of language processes: The Sperlonga Meeting II*. Hillsdale, NJ: Lawrence Erlbaum.
- Marslen-Wilson, W., & Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.
- Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, 43, 240-274.
- Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for a model of lexical retrieval. *Journal of Memory and Language*, 28, 462-485.
- McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII*. Hillsdale, NJ: Lawrence Erlbaum.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: 1. An account of basic findings. *Psychological Review*, 88, 375-407.

- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-466.
- Meyer, A. S. (1991). The timecourse of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 30, 69-89.
- Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, 20, 715-726.
- Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. *Journal of Memory and Language*, 24, 3-26.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence for a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory & Cognition*, 2, 309-321.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- O'Seaghdha, P. G., Dell, G. S., Peterson, R. R., & Juliano, C. (1992). Models of form-related priming in comprehension and production. In R. G. Reilly & N. E. Sharkey (Eds.), *Connectionist approaches to natural language processing* (pp. 373-408). Hillsdale, NJ: Lawrence Erlbaum.
- Perfect, T. J., & Hanley, J. R. (1992). The tip-of-the-tongue phenomenon: Do experimenter-presented interlopers have any effect? *Cognition*, 45, 55-75.
- Peterson, R. R., Dell, G. S., & O'Seaghdha, P. G. (1989). A connectionist model of form-related priming effects. *Proceedings of the 11th Annual Conference of the Cognitive Science Society* (pp. 196-203). Hillsdale, NJ: Lawrence Erlbaum.
- Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 43-102.
- Ratcliff, F., & Hartline, H. K. (1959). The response of Limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *Journal of General Physiology*, 42, 1241-1255.
- Roediger, H. L., & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology*, 36, 213-242.
- Roediger, H. L., Neely, J. H., & Blaxton, T. A. (1983). Inhibition from related primes in semantic memory retrieval: A reappraisal of Brown's (1979) paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 478-485.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.
- Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89, 60-94.
- Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986a). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.
- Rumelhart, D. E., & McClelland, J. L. (1986b). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 216-271). Cambridge, MA: Bradford Books.

- Rumelhart, D. E., & Norman, D. A. (1982). Stimulating a skilled typist: A study of skilled cognitive motor performance. *Cognitive Science*, 6, 1-36.
- Schade, U. (1992). *Konnektionismus - Zur Modellierung der Sprachproduktion*. Opladen: Westdeutscher Verlag.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the timecourse of lexical access in production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 65-76.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of visual word recognition and naming. *Psychological Review*, 96, 523-568.
- Sevold, C., & Dell, G. S. (submitted). *Planning units in speech production: Evidence for sequential activation of units inside the syllable*.
- Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 1239-1250.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143-186). Hillsdale, NJ: Lawrence Erlbaum.
- Stemberger, J. P. (1990). Wordshape errors in language production. *Cognition*, 35, 123-158.
- Stemberger, J. P. (1991). Apparent antifrequency effects in language production: The addition bias and phonological underspecification. *Journal of Memory and Language*, 30, 161-185.
- St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46, 217-257.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 545-569.
- Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. A. (1979). Evidence for multiple stages in the process of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 18, 427-441.
- Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language*, 26, 608-631.
- Wheeldon, L. (1989). *Priming of spoken word production*, Unpublished doctoral dissertation, University of Cambridge.
- Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.