Psychological Review
1986, Vol. 93, No. 3, 283-321


0033-295X/86/$00.75

# A Spreading-Activation Theory of Retrieval in Sentence Production

Gary S. Dell
University of Rochester

This article presents a theory of sentence production that accounts for facts about speech errors—the kinds of errors that occur, the constraints on their form, and the conditions that precipitate them. The theory combines a spreading-activation retrieval mechanism with assumptions regarding linguistic units and rules. Two simulation models are presented to illustrate how the theory applies to phonological encoding processes. One was designed to produce the basic kinds of phonological errors and their relative frequencies of occurrence. The second was used to fit data from an experimental technique designed to create these errors under controlled conditions.



Psycholinguistics is concerned with three basic and interrelated aspects of language—acquisition, or how language is learned; comprehension, or how sentences are understood; and production, or how sentences are spoken. Of these three, production has been the least studied and is probably the least understood.

Recent research, however, has established production as an equal partner in the psycholinguistic trinity by defining problems that are unique to it. For example, there is the problem of how thought, which often lacks a temporal structure, is translated via syntactic rules into temporally ordered speech (e.g., Bock, 1982; Kempen, 1978; Lapointe, 1985; Levelt, 1982; MacKay, 1982), or the problem of editing (e.g., Levelt, 1983; Motley, Camden, & Baars, 1982), or that of specifying how language production fits into a general theory of action (Fowler, Rubin, Remez, & Turvey, 1980; Norman, 1981). At the same time new methodologies for studying production have been developed, and production is now routinely studied in controlled experiments, as well as through naturalistic observation (e.g., Bock, 1983; Cooper & Paccia-Cooper, 1980; Fromkin, 1980; Sternberg, Monsell, Knoll, & Wright, 1978).

This article presents a new theory of production. Although current theories have been broad (e.g., Bock, 1982; Butterworth, 1981; Garrett, 1975; Shattuck-Hufnagel, 1979; Stemberger, 1982a) and formal (e.g., MacKay, 1982) they have not been developed to the point that a large number of precise quantitative predictions can be made and tested—to the point that data can be fitted to models. The theory presented here is a step toward reaching this goal.

Specifically, I will deal with what Garrett (1975) has called sentence production, which can be defined as the processes by which a semantic representation of a sentence-to-be-spoken is translated into a phonetic representation that can guide the articulatory musculature and produce speech. So, it is not a theory of why a speaker says what is said but of how it is said.

During the translation from meaning to sound, words must be selected and ordered in adherence with the rules of the grammar of the speaker's language. This I call syntactic encoding. These words must be specified in terms of their constituent morphemes (morphological encoding), and these morphemes must be spelled out in terms of their sound (phonological encoding). Although the theory deals informally with all of these encoding processes, I will focus on phonological encoding, and particularly the retrieval of phonological forms, for the theory's formal development. This is not because phonological encoding is the most basic of these three processes, but because it is the most tractable. The units and rules of phonology are both simpler and much less numerous than those of the other processes, and so it is easier for an extensive formal model to lead to quantitative predictions.

The theory combines assumptions from linguistic theory regarding the units and structures of language with a precisely specified retrieval mechanism based on *spreading activation.* As such, the theory falls into a general class of theories that includes interactive activation models (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Stemberger, 1982a), connectionist models (Cottrell & Small, 1983; Feldman & Ballard, 1982; Grossberg, 1982; Grossberg & Stone, 1986), and other spreading activation theories (e.g., Anderson, 1976, 1983; Dosher, 1982; Rumelhart & Norman, 1982). Although these kinds of theories have their critics (e.g., Ratcliff & McKoon, 1981), they are, in many ways, beginning to form a theoretical paradigm in cognitive psychology. The most direct ancestors of the theory are spreading-activation production models that postulate a network of linguistic rules and units in which decisions about what unit or rule to choose are based on the activation levels of the nodes representing those rules or units (Bock, 1982; Crompton, 1981; Dell & Reich, 1977, 1980, 1981; MacKay, 1982; Roche-Lecours & L'hermitte, 1969; Stemberger, 1982a). Another important antecedent of the theory is found in frame-and-slot models of production, in which it is proposed that internal representations of utterances-to-be-spoken are constructed by inserting linguistic items into slots

This work was supported by National Science Foundation Grant BNS-8406886 and National Institute of Child Health and Human Development Grant HD-13318.

The author wishes to thank Thomas Berg, Tom Bever, William E. Cooper, Joe Stemberger, Carol Fowler, Susan Garnsey, Steve Lapointe, Gail McKoon, Roger Ratcliff, Peter Reich, Mike Tanenhaus, and three anonymous reviewers for their helpful comments.

Correspondence concerning this article should be sent to Gary S. Dell, Department of Psychology, University of Rochester, Rochester, New York 14627.

in independently created structures that define the order of the slots (Garrett, 1975; MacKay, 1972; Reich, 1977; Shattuck-Hufnagel, 1979).

The data to be accounted for by the theory are facts about speech errors, or slips of the tongue. These errors, often called lapses, Spoonerisms, or Freudian slips, have long been cited as the best available data for the study of language production. Freud (1901/1958) argued that slips, in addition to revealing unconscious motives, may provide "insight into the probable laws of the formation of speech" (p. 71). Meringer and Mayer (1895), who were among the first to draw attention to slips as a data source, hoped to discover the "mental mechanism in which the sounds of a word or sentence . . . are linked" (p. 10). More recently, Fromkin (1973) remarked that "speech error data . . . provide us with a window into linguistic processes and provide, to some extent, the laboratory data needed in linguistics" (pp. 43–44). Each of these researchers recognized that the inner workings of a highly complex system are often revealed by the way in which the system breaks down. The theory presented here specifies the kinds of breakdowns that are likely to occur, the constraints on the form of slips, and the conditions that precipitate them. In addition, the theory will deal with the tendency for errors to create meaningful combinations of items, the tendency for similar items in similar environments to slip with one another, and the influence of changes in the speaking rate on errors.

The article consists of four sections. The first section presents some of the basic facts about speech errors and the methodology associated with their study. The second section is a general description of the theory as it applies to syntactic, morphological, and phonological encoding. In the third section, the theory is formally developed with respect to phonological encoding. A simulation model of the retrieval and ordering of speech sounds is presented to demonstrate how the theory accounts for error phenomena. In addition, a model of a particular experimental paradigm for generating phonological slips is developed and used to fit data from a comprehensive experiment using this paradigm. The final section briefly suggests how the model can be extended to deal with morphological and syntactic encoding.

## Slips of the Tongue

A slip of the tongue can be defined as an unintended, nonhabitual deviation from a speech plan. By requiring that a slip be unintended, the definition eliminates witticisms and normal phonological deletions and assimilations (e.g., *Did you eat yet?* being spoken as *Jeet chet?*). Eliminating nonhabitual errors removes from consideration nonstandard expressions (*between you and I*), classical malapropisms (consistently saying *reminisce* for *remiss*), and errors that arise from pathology. Although habitual errors are clearly related to normal slips, they are not considered here.

Most of what is known about slips comes from analyses of collections of errors (e.g., Boomer & Laver, 1968; Fay & Cutler, 1977; Fromkin, 1971; Garrett, 1975; McKay, 1970; Nooteboom, 1969; Shattuck-Hufnagel, 1979; Stemberger, 1982a). These collections were composed of errors that were personally heard and noted by the investigators. Although there are many problems associated with this technique of gathering data (see Cutler, 1981), the analyses have been useful in three ways. First, they have identified the types of slips that occur and some of the constraints on their form. Just knowing which slips are possible and which are not severely constrains theories of production. Second, there are many regularities in error collections, some of which are so obvious as to rule out the possibility that they arose from the sampling biases associated with the collection procedures. These regularities must be accounted for, as well. Finally, the analyses of natural errors have led to the development of techniques for creating errors in controlled experimental settings (e.g., Baars & Motley, 1974). The natural errors have suggested both how errors might be elicited and some potentially interesting independent variables.

Table 1 summarizes some of the basic types of speech errors. The first aspect of slips that one notices is that many different-sized linguistic units can slip. Usually slips are divided into sound, morpheme, and word errors depending on the size of the unit disrupted. Sound errors can be further divided into errors involving single speech sounds (phonemes), consonant clusters, and vowel–consonant combinations. Most vowel–consonant errors involve the rime constituent (vowel plus remaining consonants in the syllable), and not the whole syllable. In addition, errors that can be analyzed as phonemic feature slips have been found (Fromkin, 1971), but these appear to be rare (Shattuck-Hufnagel & Klatt, 1979).

In addition to the involvement of different linguistic units, errors differ as to the nature of the disruption. On the one hand are errors that involve some kind of misordering, sometimes called *syntagmatic* or *contextual errors.* In this category are exchanges, anticipations, perseverations, shifts, anticipatory and perseveratory additions, and some kinds of deletions. On the other hand are errors for which it is difficult to identify a source within the utterance. These are the *noncontextual errors,* a category that includes noncontextual substitutions, blends, additions, and deletions.

Exchanges, anticipations, perseverations, blends, and noncontextual substitution errors can be characterized by two units, which are referred to as the *interacting units.* For exchanges, the two exchanging units are the interacting units; for anticipations and perservations, the target, or intended unit, and the unit that moves to replace it are the two interacting units; and for substitutions, the target unit and the one replacing it are the interacting units.

Most of the quantitative facts about slips refer to relations that often hold between the two interacting units, such as their dimensions of similarity and difference, or in the case of misordering errors, the distance between the units. For example, one such relation that is often noted is that the interacting units are nearly always, or at least very often, members of the same linguistic category. In the word errors presented in Table 1, one can see that the interacting words are members of the same syntactic category. Sound and morphemic errors also show a categorical constraint—the relevant categories are not syntactic ones, though. In sound errors, vowels interact with other vowels, initial consonants with other initial consonants, and final consonants with other final consonants. At the morphological level, stems interact with stems, and affixes with affixes.

## Introduction to the Theory

The theory has its roots in an assumption regarding the production of ordered behavior that can be attributed to Lashley

Table 1
*Types of Speech Error*

| Type | Examples | Unit involved[a] |
|---|---|---|
| Sound errors | | |
| Misordering | | |
| Substitution | | |
| Exchange | *York library → lork yibrary* | Phoneme |
| | *Spill beer → speer bill* | Rime constituent |
| | *Snow flurries → flow snurries* | Consonant cluster |
| | *Clear blue → glear plue* | Feature |
| Anticipation | *Reading list → leading list* | Phoneme |
| | *Couch is comfortable → comf is . . .* | Syllable or rime |
| Perseveration | *Beef noodle → beef needle* | Phoneme |
| Addition | | |
| Anticipatory addition | *Eerie stamp → steerie stamp* | Consonant cluster |
| Perseveratory addition | *Blue bug → blue blug* | Phoneme |
| Shift | *Black boxes → back bloxes* | Phoneme |
| Deletion[b] | *Same state → same sate* | Phoneme |
| Noncontextual errors | *Department → jepartment* | Phoneme |
| (substitution, | *Winning → winnding* | Phoneme |
| addition, | *Tremendously → tremenly* | Syllable |
| deletion) | | |
| Morpheme errors | | |
| Misordering | | |
| Substitution | | |
| Exchange | *Self-destruct instruction → self-instruct de . . .* | Prefix |
| | *Thinly sliced → slicely thinned* | Stem |
| Anticipation | *My car towed → my tow towed* | Stem |
| Perseveration | *Explain . . . rule insertion → . . . rule exsertion* | Prefix |
| Shift | *Gets it → get its* | Inflectional suffix |
| Addition | *Dollars deductible → dedollars deductible* | Prefix |
| | *Some weeks → somes weeks* | Inflectional suffix |
| Noncontextual errors | *Conclusion → concludement* | Derivational suffix |
| (substitution, | *To strain it → to strained it* | Inflectional suffix |
| addition, | *He relaxes → he relax* | Inflectional suffix |
| deletion) | | |
| Word errors | | |
| Misorderings | | |
| Substitution | | |
| Exchange | *Writing a letter to my mother → writing a mother to my letter* | Noun |
| Anticipation | *Sun is in the sky → sky is in the sky* | Noun |
| Perseveration | *Class will be about discussing the test → . . . discussing the class* | Noun |
| Addition | *These flowers are purple → these purple flowers are purple* | Adjective |
| Shift | *Something to tell you all → something all to tell you* | Quantifier |
| Noncontextual errors | | |
| Substitution | *Pass the pepper → pass the salt* | Noun |
| | *Liszt's second Hungarian rhapsody → second Hungarian restaurant* | Noun |
| Blend | *Athlete/player → athler* | Noun |
| | *Taxi/cab → tab* | Noun |
| Addition | *The only thing I can do → the only one thing* | Quantifier |
| Deletion | *I just wanted to ask that → I just wanted to that* | Verb |

*Note.* These errors are taken from Stemberger (1982a), Fromkin (1971), Garrett (1975), Shattuck-Hufnagel (1979), and Dell and Reich (1981). In the examples, the text to the left of the error is the intended utterance and the text to the right is the actual utterance.

[a] The units involved in the examples for a given type do not exhaust the set of possible units for that type. In general, a given type can occur with many different units.

[b] The deletion category appears both under the heading of sound misorderings and sound noncontextual errors. This is because some deletions, such as the *same state* example seem to involve a contextual influence, whereas other deletions do not.

(1951), and was originally applied to sentence production by Fry (1969). This assumption is that production involves the construction of at least one internal representation of the planned behavior and that this representation is built up in advance of the overt behavior that it represents. Lashley referred to production units that have been planned in advance as "primed," and suggested that this phenomenon might be responsible for ordering errors. In applying this notion to sentence production, Fry identified five internal representations corresponding to semantic, syntactic, morphological, phonological, and motor encoding. Higher order levels, such as the semantic representation can be assembled before the lower order ones, thus necessitating the need for the storage or buffering of the advance planning. Reich (1977) showed how slips of the tongue could result from the need for storing advance planning at the higher levels in a buffer. When the contents of the buffer

are addressed in order to construct the lower representation, errors result because of the co-occurrence of more than one unit in the buffer. The idea that slips result from such a co-occurrence, or *computational simultaneity* (Garrett, 1975), of representational units has been either an implicit or explicit assumption of all recently proposed accounts of production, and the theory presented here continues in that tradition.

The theory brings together aspects of linguistic competence and a set of processing assumptions that have often appeared in current information processing models. I hope to show that this wedding of linguistic and psychological theory leads naturally to predictions regarding speech errors. First, I will outline some of the premises of linguistic theory that are important, and then describe the processing assumptions.



## *Linguistic Assumptions*

Two very general linguistic distinctions are important to the theory—the distinction among linguistic levels and the distinction between generative rules and the lexicon.

Current linguistic theory recognizes that a speaker's knowledge of his or her language can be organized into separate components, or levels. Typically, semantic, syntactic, and phonological levels are proposed, and often a morphological level concerned with word formation is distinguished, as well. The motivation for this separation stems from the fact that language is productive; that is, language users can produce and understand novel utterances. Specifically, there are qualitatively distinct classes of productivity: syntactic productivity, which reflects the potential to combine words in novel ways; morphological productivity, illustrated by the speaker's ability to create new words out of existing morphemes (e.g., *my musical tastes are pre-Bachian*); and phonological productivity, which can be seen in the speaker's recognition that certain combinations of sounds, although nonwords, are nonetheless well-formed sequences (e.g., *snurk*). Each type of productivity can be described by generative rules (syntactic, morphological, and phonological rules) that tell us whether a given combination of units is a potential string in the language. In the sense that these sets of rules are very different from one another—they operate on differently sized units and are conditioned by different factors—language can be said to have levels.

Different linguistic theories have somewhat different conceptions about linguistic levels, particularly the syntactic level. For example, the transformational generative approach (e.g., Chomsky, 1981) divides the syntax into subcomponents (e.g., base rules, movement rules) associated with different structural descriptions of the sentence (e.g., D-structure, S-structure, surface structure). I use the term *syntactic level* to refer to the whole syntactic system, without a specific commitment to this substructure. However, when I refer to the *syntactic representation* in the production theory, I refer to a representation that is the final product of syntactic processes, a surface structure of the sentence in which terminal nodes are lexical items that have yet to be specified in terms of their internal morphological or phonological structure.

The notion of separate levels, each with its set of generative rules, is central to the theory of production developed here. Slips of the tongue can be seen as products of the productivity of language. A slip is an unintended novelty. Word errors create syntactic novelties; morphemic errors create novel words; and sound errors create novel, but phonologically legal, combinations of sounds.

The second important aspect of linguistic theory is the distinction between information represented as generative rules and information stored in the lexicon. Each level is associated with a set of rules that defines the combinatorial possibilities of units at that level. As stated above, these rules codify the productivity associated with each level. Rules are stated in terms of categories of unit types—syntactic rules in terms of syntactic categories (noun, verb, etc.), morphological rules in terms of morphological categories (stem, prefix, suffix, etc.), and phonological rules in terms of phonological categories (initial stop, vowel, etc.).

Unlike the generative rules, which code productive knowledge, the lexicon contains nonproductive stored knowledge. For the sentence-production theory the lexicon is represented as a network rather than a listing (e.g., Lamb, 1966; Lockwood, 1972; Reich, 1970). The network contains nodes for linguistic units such as concepts, words, morphemes, phonemes, and phonemic features, and (in some treatments) syllables and syllabic constituents as well. Conceptual nodes connect to word nodes, thereby defining the words. Words connect to morphemes, morphemes connect to phonemes, and phonemes to features. The connections specify the relation between connecting units. For example, the node for the morpheme *cat* would connect to nodes for phonemes /k/, /æ/, and /t/, with the connections further specifying that the morpheme *cat* is composed of these three phonemes in this particular order. Other componential relations are assumed to be unordered. For example, the word *cat* might connect to the conceptual features *feline* and *domesticated,* and the phoneme /k/ might connect to the phonemic features *unvoiced, velar,* and *stop,* but in both cases no particular order would be specified.

Next, I describe the nature of the interaction between the generative rules and the lexicon. Recall that the rules of each level define acceptable combinations of items at that level. A good way to think of this is to imagine that each rule system generates a frame with categorized slots. For the sentence *This cow eats grass* the syntactic rules might generate *Determiner Noun present-tense Verb Noun,* where the determiner, verb, and nouns are categorized slots. The corresponding morphological frame might be *S S SAf S* (*S* indicates a word stem and *Af* an affix). The phonological frame for *this cow* could be something like *IC V FC IC V,* where *IC, V* and *FC* are initial consonants, vowels, and final consonants, respectively. Thus, the rules define acceptability in terms of categories of items that exist in the lexicon. In order to fill in these categorically specified slots, the lexicon must be addressed. So, at each level, a lexical selection process occurs, words being selected at the syntactic level, morphemes at the morphological level, and phonemes and/or features at the phonological level. In order for this selection process to work, that is, generate a grammatical sentence, each item in the lexicon must be labeled with regard to its membership in the relevant categories. So, like in a dictionary, words must be marked as to their syntactic class. But in addition, other units that are selected by other rule systems must be marked. Morphological and phonological units must be labeled as to their morphological and phonological classes, respectively. I refer to the marking of lexical nodes with category information as insertion rules;

that is, these rules specify what nodes can be inserted into categorized slots in frames.

Thus, within linguistic theory, one can identify three basic types of linguistic knowledge: information stored in the lexicon, information represented as categorically specified rules, and the insertion rules, which relate the two former types. In the next section, I introduce in a general way the processing assumptions of the sentence production theory and outline how the three types of linguistic information are used in the theory.

### *Processing Assumptions in the Theory*

The principal assumption of the theory is that at each level a representation of the sentence-to-be-spoken is constructed. Thus, a planned utterance will exist at various times as a semantic, syntactic, morphological, and a phonological representation. The theory describes the construction of the latter three representations, that is, it describes syntactic, morphological, and phonological encoding.

First of all, what are these representations like? According to the theory, each is an ordered set of items found in the lexicon—the syntactic, morphological, and phonological representations are ordered sets of various words, morphemes, and phonemes, respectively. One way to imagine the representation at a particular level is as a collection of order tags that are attached to nodes in the lexical network, dictating the contents of the representation and their order. The processing assumptions of the theory tell how these order tags are handed out.

The construction of a representation at each level goes on simultaneously with that of the other levels, with the rate of processing depending on factors intrinsic to the level and on the rate of processing of the level immediately above it. In general, the selection of items for a lower representation must await the construction of corresponding structures of the higher representation. Obviously, a word filling a certain syntactic role must be selected before it can be phonologically encoded. However, processing at the higher level can lead that of the lower level to a variable extent. The representation, or set of tagged nodes, stores the processing of each level that has not yet been dealt with by lower levels. Thus, each representation acts as a buffer and creates what Reich (1977) has called a *loosely yoked* production system. The processing at each level is yoked to the one above it, but the higher one can get ahead.

The idea that the construction of each representation is guided by the representation above it seems reasonable, but it does raise the question of what guides the construction of the highest representation, here characterized as the semantic representation. The question amounts to that of how we decide what to say—what kind of speech act we wish to make and what propositional content we need to include. Bock (1982) termed this the *interfacing representation*—that which interfaces between language and thought. Although some psycholinguistic studies have investigated these representations in production (e.g., Levelt & Maassen, 1981), much of the significant theoretical work has come from the field of artificial intelligence (e.g., Allen, 1979). These issues are beyond the scope of this article. However, in order to say something about syntactic encoding, the theory must make some simple assumptions about the nature of the semantic representation. These are, first, that its units are concepts, which appear in the lexical network connected to word nodes; second, that these nodes behave as do other nodes in the processing assumptions of the theory; and finally, that there exists some (unspecified) mechanism that makes sure that these concepts become available at appropriate times during syntactic encoding.

The basic idea of the theory is that the tagged nodes constituting a higher representation activate nodes that may be used for the immediately lower representation through a spreading-activation mechanism. The lower representation is constructed as the generative rules associated with that level build a frame, or ordered set of categories, and the insertion rules fill in the slots of the frame through a decision rule based on the activation levels of the nodes representing candidate items. When an item is selected for a slot, it becomes part of the developing lower representation, and so it receives its tag. Thus, a principal mechanism for the translation of information from one representation to another is spreading activation through the lexicon.

*Processing in the lexicon: Spreading activation.* In this section I describe the retrieval processes in the lexical network. These processes cause the units appropriate for a given representation to become activated. The mechanism is that of spreading activation. This aspect of the present theory is most important, in that just about all of its substantive predictions regarding speech errors derive from spreading activation, or from interactions between spreading activation and other assumptions.

Unlike many spreading-activation theories, in which activation is assumed to be a binary variable, the present theory assumes that activation level is a real variable, with the activation level of node $j$ at time $t$ given by $A(j, t)$. Although the theory is not incompatible with negative activation levels, only positive values occur in its present conception. The defining components of spreading activation are *spreading, summation,* and *decay.* When a node has an activation level greater than zero, it sends some proportion of its activation level to all nodes connected to it (spreading). This proportion is not necessarily the same for each connection. When the activation that is sent out reaches its destination node, it adds to that node's current activation level (summation). For the sake of simplicity I assume no thresholds, saturation points, or other nonlinearities in the spreading process. It is necessary, however, to include a passive decay of activation over time to keep levels down. Specifically, activation is assumed to decay exponentially toward zero. These operations, spreading, summation, and decay, apply to all of the nodes in the lexical network at all times, regardless of whether the node is part of a representation (tagged or not).

In order to describe the spreading process more formally, I use the concept of quantized time—that is, that time is assumed to pass in discrete steps. The activation levels of all of the nodes at time $t_1$ provide the basis for calculating the new activation levels for the nodes at $t_2$, which in turn, are used to compute the levels at $t_3$, and so on. The rules of spreading activation are given by

$$A(j, t_i) = [A(j, t_{i-1}) + \sum_{k=1}^{n} p_k A(c_k, t_{i-1})](1 - q),$$

where $A(j, t_i)$ is the activation level of node $j$ at a particular time; $t_i$, $A(j, t_{i-1})$ is its activation level at the preceding time step; $c_1$, $c_2$, $c_3 \cdots c_n$ are all nodes directly connected to $j$; $p_1, p_2, p_3, \cdots$

$p_n$ are the spreading rates associated with the connections from $c_1, c_2, c_3 \cdots c_n$ to node $j$; and $q(0 < q < 1)$ is the decay rate. The spreading rates are assumed to be greater than zero. Hence, the theory emphasizes excitatory rather than inhibitory processes. I should point out that the behavior of the activation levels (whether they grow without bound or diminish over time) depends on the values of the $p$s, the value of $q$, and the structure of the network. In the specific models of the spreading process, which are presented in the section on phonological encoding, the parameters are set so that, given a particular network, activation does not grow without bound.

One of the important assumptions regarding spreading activation in the theory is that all connections are two way. If Node A connects to B, then B connects to A. Given the nature of the connections and the hierarchical structure of the lexical network, each connection can be classified as either excitatory top-down, or excitatory bottom-up. For each top-down connection, such as that from a particular morpheme to a particular phoneme, there is a bottom-up connection in the reverse direction. These bottom-up connections deliver positive feedback from later to earlier levels and play a critical role in the theory. Their presence makes processing in the network highly interactive and, as will be seen, generates some nonobvious predictions.

The spreading-activation assumptions only specify how activated nodes affect one another. Clearly, more is needed. The pattern of activation among the nodes must be translated into a representation. That is, there must be a way of deciding which activated nodes get tagged as part of the representation. This function is carried out by the categorical rules, which build the representation's frame, and the insertion rules, which specify what nodes can fill particular slots in the frame.

*Constructing representations.* In this section I outline how a lower representation is constructed given a higher representation. The first important concept is that of the *current* node. It is that item of the higher level representation that is in the process of being translated into corresponding items at the immediately lower level. Because processing can occur on more than one level, there would typically be more than one current node, one for each level's representation that has not been fully translated into a lower level. When the construction of a lower representation begins, the current node is that node of the higher representation that is tagged as first. The initial step in the translation is the activation of the current node. Its activation level is increased by some arbitrary amount called *signaling activation.* This increase in activation will then make itself felt as increases in activation in nodes connected to the current node through spreading, as I have outlined. While this is going on, the rules associated with the lower level are constructing a frame of categorized slots. Other than to say that the operation of these rules is guided by input from both the higher and lower representations, I leave this issue for development in the particular sections on phonological, morphological, and syntactic encoding. After a categorically defined slot is created in the frame, the insertion rules operate to fill that slot. They look at the activation levels of all nodes that are marked with the specified category and select that item whose node possesses the highest level. The selection of an item is followed immediately by the reduction of its activation level to zero (to prevent its being selected over and over again) and the tagging of the node (its order having been defined by the frame). A final action that may follow the selection of an item is that of changing the higher representation's current node. When this occurs depends on the nature of the mapping from the higher to the lower representation. If an updating of the current node is called for, the node whose tag places it immediately after the previous current node becomes the current node.

A representation can have a given node tagged more than once. In fact this is necessary when items are repeated. For example, in the sentence *people need people* the syntactic representation would have the word *people* tagged twice; the phonological representation would have /i/ tagged three times. A general feature of this model, and related models (e.g., MacKay, 1982; Stemberger, 1982a) is that representations are not built up of tokens of items, but instead consist of instructions to use (and reuse, if necessary) item types. Whereas token-based models might have two tokens of *people* to represent the above sentence, a type-only model would just have one that is used twice.

Although we will see later how the type-only nature of the theory's representations enables it to explain some repeated-item effects in errors, it does introduce one problem. Because on occasion you need to tag a node that is already tagged, you can run into problems when representations on two adjacent levels are both being constructed. Consider what happens as the syntactic representation for *people need food* is constructed. After *people* and *need* have been selected and tagged, the syntactic rule system is seeking a noun, and in this case *food* should be the most highly activated noun. However, if we assume that the next lower representation (here assumed to be the morphological representation) is beginning to be constructed as well, then the word node for *people* will also be highly activated as it becomes the current node—meaning that node that is currently being translated into units of the lower representation. If *people* becomes current just when the syntax is looking for the noun *food,* then *people,* being a noun as well, may be selected again, replacing *food.* The problem is essentially one of confusion between levels. The activated current node *people* is current with respect to the lower morphological level. Its activated state is irrelevant to the continued construction of the syntactic representation that is proceeding two words down the line, at *food.* Yet the syntactic insertion rule that is looking for the noun following the verb *need,* doesn't know this. It only looks at the activation levels of all of the word nodes in the appropriate category, and *people* may be the most activated one.

A radical solution to this problem is to tightly control the timing of representation building. For example, one could prevent the simultaneous construction of adjacent levels by stipulating that no level with unfilled frame slots can have a current node. I find this unappealing for two reasons. First, current information-processing theory dealing with stimulus recognition has emphasized a loosely yoked, or cascaded, view of timing between levels (e.g., McClelland, 1979). It is reasonable to presume that word production works similarly. Second, as will be seen, the production theory explains anticipatory speech errors at a given level by appealing to ongoing processing at the next highest level. If that level is quiescent, such slips are unexplained. A solution that maintains the loosely yoked hierarchy of representations is to allow a current node to be reselected, but only if its activation level is significantly higher than its competitors in the same category. Although I do not develop this

latter solution in any more detail, I will tentatively adopt it for the theory until a less ad hoc solution emerges.

A final aspect of the building of representations in the theory is the role of the speaking rate. I assume that the speaking rate is, to some extent, under the speaker's control. When the speaker selects a rate, he or she is essentially controlling the rate of construction of the representations, specifically the rate that categorized slots become available and are filled. The theory is neutral on the question of whether rate control affects the rate of building the frames or the rate with which slots are filled. The higher level representations need not be constructed exactly in step with articulation because of the loosely yoked nature of the system. However, lower representations, such as the phonological representation, are assumed to be constructed in much the same time frame as that of articulation. The spreading and decay rates are assumed not to vary with the speaking rate. As a result, the time needed for higher level items to retrieve lower level items is a limiting factor in the speaker's ability to produce correct speech, according to the theory.

*Production of a sentence.* By putting together the theory's assumptions, one gets a general picture, such as that in Figure 1, which presents a moment in the production of the sentence *Some swimmers sink.* In this particular example, I assume that the word *swimmer* has two morphemes, but that the word is also present in the lexicon as a noun. Also, the morphemes for plural and other inflections are generated by the syntax and attached to the stems in the morphological representation. Finally, it is assumed that the phonological rules employ the categories initial-consonant group (onset), vowel (nucleus), and final-consonant group (coda). These particular assumptions are put forward just for the example, in order to illustrate the general properties of the theory.

First, consider the syntactic level. Here the representation consists of word nodes (or what Kempen & Huijbers, 1983, call *lemmas*) that are linked to slots in a syntactic frame that codes surface phrasal structure. In the figure the frame of categories is being built, and two of its slots, the quantifier and noun slots, have already been filled in, as indicated by the order tags on word nodes *some* and *swimmer.* In addition, a node for the plural has received an order tag. The insertion rule for verbs is presently searching out that verb with the highest activation level for inclusion in the representation. Because the verb *sink* has the highest activation level of the verbs (indicated by highlighting), it will be selected and receive the next order tag. Notice that other word nodes are activated at this time. Words that have already been selected for inclusion in the frame remain somewhat activated, despite the rule that the activation level of selected items is set to zero. This is because when these nodes were selected they had high activation levels, and as a result their neighbors were activated as well. Upon being selected, such a node's activation level dropped to zero, but it would normally rebound from zero very quickly because of its many activated neighbors. So, although selection of a node, in principle, entails the loss of all activation, in practice, a selected node's activation rebounds and then gradually dies away. Another reason that previously selected word nodes remain activated is that some of them either are, or have been, current nodes for the construction of the lower morphological representation. This is the case with *swimmer,* which is a current node in Figure 1. (Recall that when a node is first granted current node status its activation level is boosted.) Not only are previously selected nodes activated, but also ones that represent upcoming words in the representation. Their activation comes about because the concepts that define them (and hence are connected to them) are part of the semantic representation (which is not shown in Figure 1). Other nodes that are active include some that are not, nor ever will be, part of the syntactic representation. These nodes become active simply as a consequence of activation spreading. For example, nodes that have related conceptual structures with those in the representation are activated because the semantic representation includes conceptual nodes that exist in the lexical network and participate in spreading activation. The node for *drown* might be activated in this particular case. In fact, because it is a verb it would compete with *sink* for the slot currently being filled.

Turning to the morphological level, it can be seen that the nodes are marked for morphological category, not just syntactic category. (There is considerable debate in linguistics as to just what these morphological categories are. I will return to this question later; for the purposes of this example I assume that the morphemes *sink* and *swim* belong to a stem category $S_v$, for verb stem; *some* is a quantifier stem, $S_q$; the derivational affix *er* belongs to affix category $Af_1$; and the inflectional affix for the plural belongs to $Af_2$.) In the figure the morpheme *some* has already been selected. The morphological frame for the next word, *swimmers,* has been constructed, and the insertion rule for verb stems is seeking out that stem with the highest activation level. Of the stems shown in Figure 1, it should select *swim* over *sink. Sink* is just not highly activated at this time, and *some,* although highly activated, is not of the same morphological category as *swim.* After selection of *swim* for the verb stem, the most highly activated affix in category $Af_1$ will be sought out and when selected, given a tag after that of *swim.* At this point, the current node of the syntactic representation should be changed. It has been the word node *swimmer* as shown in Figure 1. After the selection and tagging of a stem *swim* and an affix *er,* the process should change the current node of the syntactic representation from *swimmer* to plural and then the plural morpheme can be selected.

At the phonological level, no representation has yet been built. The morpheme *some,* the current node for this level, is activated and is retrieving its phonemes via spreading activation. At present the insertion rule for initial consonants is seeking out the most highly activated onset (*initial* consonant or cluster). Notice again that nodes representing phonemes for upcoming morphemes are activated, as well as those of *some.* It should be noted as well that at the phonological level items from the minor syntactic categories (closed-class items), such as *some,* are encoded in the same way as those from major syntactic categories (open-class items), such as *swim,* contrary to other proposals (e.g., Garrett, 1975).

## *Speech Errors in the Theory*

Having identified the processing assumptions of the theory, I turn to the question of where speech errors come from. Which processes are at fault when an error occurs? The theory's answer is that no one process is at fault. Errors are the natural consequences of the theory's assumptions. It just so happens that, on occasion, incorrect items have higher activation levels than

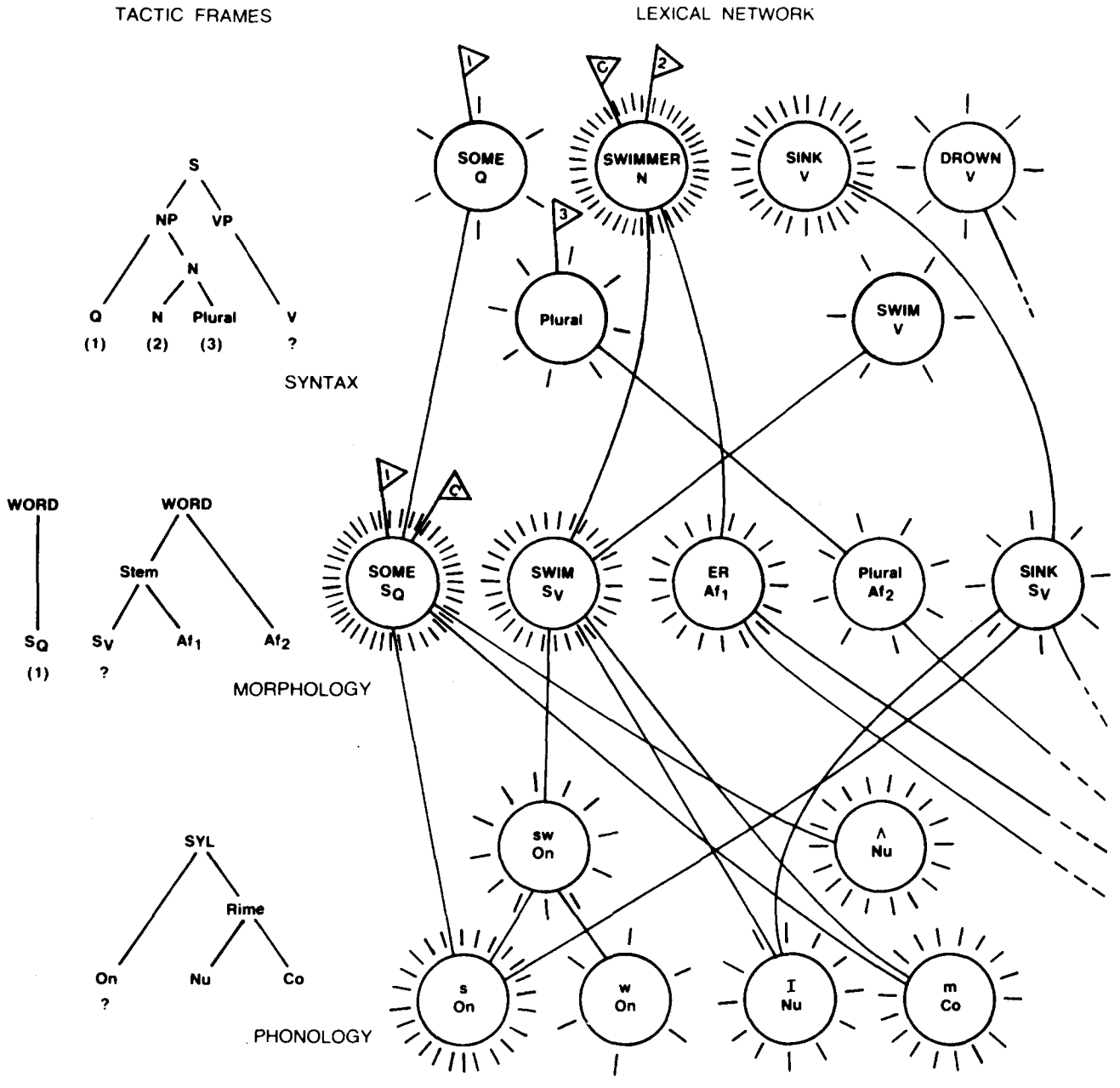


*Figure 1.* A moment in the production of the sentence "Some swimmers sink." (Tactic frames [left] specify an ordered set of categorically labeled slots. The numbered slots have already been filled in; a numbered flag (an order tag) has been placed on each node in the lexical network that stands for an item filling a slot. The question mark indicates the slot in each frame that is presently being filled. The highlighting surrounding the nodes reflects activation level and the flag with a "C" on it marks the current node of each level. Each node is labeled for membership in some category: Syntactic categories for words are (Q)uantifier, (N)oun, (V)erb, plural marker; morphological categories for morphemes are (S)tem, (Af)fix; and phonological categories for sounds are (On)set, (Nu)cleus, (Co)da. Many nodes have been left out to simplify the network, including nodes for syllables, syllabic constituents, and features.)

correct ones and are selected. A large part of this potential for error comes from the spreading of activation and the construction of multiple representations. This source of variability in activation levels results directly from the theory's assumptions. The theory does, however, admit the fact that the production of a sentence takes place amidst a background of other sources of verbal activation, which would be specified in a complete theory of language production and comprehension. For example, in the planning of an utterance many concepts would legitimately become activated that would not actually appear in the utter-

Table 2
*Some Possible Errors From "Some Swimmers Sink"*

| Error | Type |
|---|---|
| *Sim swimmers sink* | Phoneme anticipation |
| *Swum simmers sink* | Phoneme shift |
| *Some simmers sink* | Phoneme deletion |
| *Sim swummers sink* | Phoneme exchange |
| *Some swummers sink* | Phoneme perseveration |
| *Some swinkers sink* | Cluster anticipation |
| *Some sinkers swim* | Stem exchange |
| *Some swimmers swim* | Stem perseveration or word substitution |
| *Some swimmers drown* | Word substitution |

*Note.* All of these errors assume the correct construction of frames. Errors in which the frames are wrong (e.g., *Somes swimmers sink, Some swim sinkers*) are considered in the section on morphological and syntactic encoding.



ance. This background activation might include activation from concepts that were either presuppositions or inferences that were necessary in the semantic and pragmatic planning of the utterance. For example, if one were to say *Could you close the door?,* one would certainly have processed the presupposition that the door was open. As a result the concept for *open* might be active, and because of the spreading of activation, the word, morpheme, and phoneme nodes associated with the concept *open* would become activated, perhaps resulting in the slip *Could you open, I mean close. . . .* In this way, the background activation of the lexical network provides a source of variability in the patterns of activation that result during production. In the simulation models of phonological encoding presented in a later section, I introduce random fluctuations in activation levels of some nodes to model this background. However, this variability is not seen as resulting from some sort of randomness at the neural level, but rather from the multitude of processes associated with saying a sentence.

For the remainder of this section I will introduce the speech-error phenomena accounted for or predicted by the theory. The theory makes claims about the kinds of errors that will be made, the constraints on these errors, and some of the factors that contribute to their occurrence. Table 2 summarizes some of the error types that one might expect in the theory, using the sentence *Some swimmers sink* as a target; I refer to some of these in the following discussion.

To understand speech-error phenomena in the theory, one must first get a feel for how order information is represented. A good deal of the ordering is accomplished by the rules that build the frame, the ordered set of categorically specified slots. So part of the answer as to which item goes before which is given by the knowledge that certain categories of items come before other categories of items. In the present theory, however, both an item's category and its activation level interact to determine its serial position in a representation. For example, the syntactic frame for the sentence *Cows eat grass* (*N pl V N*) tells us that the nouns surround the verb. It does not tell us which noun is first. According to the theory, the decision as to which noun to put first is resolved by activation levels. At the time of selecting the first noun, *cow* should be more activated than *grass,* and the reverse should be true when selecting the second noun. Note that these differences in activation levels have to be determined by the thematic roles of *cow* and *grass* in the semantic representation, a process that the theory makes no specific claims about.

From this sort of ordering mechanism we would expect errors to follow categorical constraints. As I mentioned earlier, an error occurs when a wrong item is more activated than the correct one and is selected and tagged. However, for this wrong item to be selected it must be a member of the same category as the correct item. The theory assumes that the insertion rules, whose job it is to pick out the most highly activated node of a specified category, stick to their categories. The consequence is a strict categorical constraint on errors. For example, when a given syntactic insertion rule selects the "wrong" word, that word must, nevertheless, be of the same syntactic category as the one that belongs in that slot. Thus, in the error *Some swimmers drown, drown* must be a verb like the correct item *sink.* In the case of a word-misordering error such as *Lifeguards save swimmers* being spoken as *Swimmers save lifeguards,* the misordered words must be of the same syntactic category. Otherwise they would not "fit" in each other's syntactic slots. The same type of constraint operates with morpheme and sound errors, except that the categories are different. In morpheme errors, certain stems can replace other stems and affixes other affixes. For sound errors, which occur at the phonological level, initial consonants replace other initials, vowels replace vowels, and so on.

The categorical constraint aspect of the theory is not new. It has precedents in many other theoretical discussions of speech errors (e.g., Bock, 1982; Dell & Reich, 1981; Fay & Cutler, 1977; Fromkin, 1971; MacKay, 1972, 1982; Reich, 1977; Shattuck-Hufnagel, 1979; Stemberger, 1982a). Although there are some differences in how categorical constraints are realized in these theories, there is a common aspect: The actual error mechanism must be separable from the part of the production process that uses categorical rules—usually the tree- or frame-building process. So the categories are not wrong, just the actual linguistic items that are associated with them.

One theory that emphasizes categorical constraints is that of Garrett (1975, 1976, 1980a, 1980b, 1982). Garrett noted that in most word-level errors (e.g., exchanges, substitutions, and blends), the interacting words belong to the same syntactic class. However, errors involving the exchange of stems or individual speech sounds do not show this syntactic constraint. For example, in the morphemic error *some sinkers swim,* the two exchanged stems come from words of differing syntactic categories.[1] From these and other aspects of errors, Garrett concluded that most word errors, which obey syntactic constraints, occur at a *functional* level of processing, whereas morpheme and sound exchanges occur at a *positional* level. At the functional level those items that are computationally simultaneous; that is, items that could interact in an error, are those that belong to the same syntactic class. In contrast, at the positional level, items are computationally simultaneous with those that are nearby in terms of the serial order of the items. Syntactic class no longer governs which items compete with which.

The present theory incorporates Garrett's insights, albeit in a different guise. The distinction between the functional and

[1] Although the interacting stems are both verbs, the categories of the words themselves (*swimmer* and *sink*) differ. I will have more to say on the relation between syntactic and morphological categories in the final section of the article.

positional levels corresponds to that between the syntactic level, on one hand, and the morphological and phonological levels, on the other. At the syntactic level an item can only compete with those that can fill its slot in the syntactic representation. In contrast, at both the morphological and phonological levels parts of words can compete with corresponding parts of nearby words; the syntactic categories of these words do not govern which items are activated at the same time, as much as their positions in the sentence do.

Within the context of the theory one can expect errors of two distinct types occurring at each level of representation—misordering or contextual errors, in which intruding item or items come from the intended utterance, and noncontextual errors, in which the interference comes from outside of the utterance. As I mentioned earlier, misorderings include anticipations, perservations, exchanges, shifts, and some types of deletion and addition errors. I will briefly introduce how the first three of these types could arise according to the theory. Anticipations occur when an upcoming item of the same category as the correct item possesses a higher activation level than the correct item. The reasons for this are varied, but chief among them is the fact that upcoming items receive activation from nodes in the higher level representation. So, in the preceding example, when the syllable /sʌm/ is encoded at the phonological level, the node for the vowel /I/ (from *swim* and *sink*) happens to be activated. This is because the syntactic and morphological representations are working on *sink* and *swimmer,* respectively. Through spreading activation from these nodes, the node for /I/ becomes activated, and if more activated than /ʌ/, it will replace /ʌ/ when the vowel of the first syllable is selected creating the erroneous syllable /sIm/. The story is the same for perseverations, except that the interference comes from already encoded items that remain activated for reasons discussed earlier. Exchanges are more complex. As in other explanations of this error type (Reich, 1977; Shattuck-Hufnagel, 1979), exchanges result when an anticipation error triggers a following perseveration involving the item that was replaced in the anticipation. When an incorrect item creating an anticipation is selected, it is tagged as part of the representation. Upon being tagged, its activation level is set to zero. The proper item, having not been selected, remains activated, and thus it is available for selection when the next slot of the same category is filled in. Consider the exchange *Some sinkers swim,* which can be analyzed as an exchange of morphemes. First, *sink* replaces *swim,* creating an anticipation. Because it is now tagged as part of the morphological representation, the activation level of *sink* is set to zero. It will rebound from zero for reasons discussed before, but if it fails to rebound to a high level, it will not be selected for the stem slot in the next word. Instead, it is likely that *swim,* which was passed over in the present word, will be the most activated stem, and it will be selected, turning the anticipation into an exchange. Clearly, a number of factors affect the likelihood of an anticipation becoming an exchange. For example, one can see that when the speaking rate is rapid, many anticipations become exchanges because the activation level of the misselected item has not had time to rebound from zero to a sufficient degree to be selected again.

The second broad class of errors, noncontextual errors, is characterized by interference from outside the intended utterance. The processes of spreading activation activate nodes that are not part of the intended utterance. Normally, the activation levels of these nodes would not be high enough to be selected and replace some correct item. However, one can imagine that they might have received activation from the various background sources of verbal activation discussed earlier, and thus it is possible for a node outside the intended utterance to have a higher activation level than those that are part of the utterance. In such a case a substitution error would occur, such as *Some swimmers drown* for *Some swimmers sink.* As is the case with errors of misordering, any item that replaces some other item in a substitution error must be a member of the same category as the proper item.

Having discussed some error types from the theory's perspective, I now turn to some of the factors that, according to the theory, should affect the likelihood of a given error occurring. From the theory, one would expect four types of influences on the probability of an error—output biases, similarity effects, speech rate effects, and distance effects. I will discuss each of these in turn.

Output biases can be seen in the tendency for speech errors to create meaningful or frequently occurring combinations of units. One such bias is lexical bias—the tendency for sound errors to create actual words, or morphemes. Although it has long been suspected that slips tend to make words—Freud (1901/1958) gave some examples—it was first demonstrated experimentally by Baars, Motley, and MacKay (1975). Using a technique to create initial consonant exchanges in the laboratory, they found that the error rate for word pairs that created words (*lewd rip → rude lip*) was at least two times greater than that for control word pairs that did not make words when their initial consonants were exchanged (*Luke risk → ruke lisk*). In addition, some investigators (Baars & Motley, 1982; Berg, 1983; Dell & Reich, 1981; Stemberger, 1984a), but not all (e.g., Garrett, 1976), have found lexical bias in natural-error corpora. However, it is the experimental data that is of the greatest interest because a lexical bias effect in error collections could reflect the perceptual biases of the collectors (Stemberger, 1984a). A related effect is semantic bias, the tendency for sound errors to create words that are semantically related to other words in the environment. Fromkin (1971) speculated that errors such as *Art of the fugue* being spoken as *Arg of the flute* reflected some kind of semantic bias. Intuitively, the error seems to be more than the exchange of the /t/ and /g/ from *art* and *fugue* and the substitution of /l/ for the /y/ in *fugue.* The fact that the word *flute* is semantically related to the topic of music appears to play a role in the error. The semantic bias effect has also been experimentally demonstrated. Motley and Baars (1976) found that a word pair such as *get one* will slip to *wet gun* more often if the articulation of the pair is preceded by the semantically related pair *damp rifle.*

Output bias effects are of great importance to a theory of production in that they are analogous to the kinds of context effects that are so prevalent in the perception of spoken and written language. The facilitory effect on the perception of letters by a word context (e.g., Reicher, 1969) and the benefit to the recognition of a word brought about by an appropriate semantic context (e.g., Meyer & Schvaneveldt, 1971) are among the most well-known. These effects have prompted many theorists to propose models of perception in which information from many sources can be brought to bear on a particular deci-

sion. The stress has been on the interactive and parallel nature of the processing in order to account for the rapidity with which diverse sources of knowledge can be accessed and used (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978; McClelland & Rumelhart, 1981).

The theory outlined in this article accounts for output bias effects in speech, in much the same way that context effects are handled in such interactive theories of perception. This can be seen by comparing it to the interactive-activation theory of written word perception developed by McClelland and Rumelhart (1981; Rumelhart & McClelland, 1982). In their theory, context effects occur largely because of positive feedback between letters and words. A given pattern of activation among the letters that corresponds to a single word is reinforced by positive feedback from that word. The positive-feedback property of spreading activation in the production theory acts in much the same way. It molds the pattern of the activation levels at a given level so as to create errors that tend to be meaningful vis à vis higher levels. At the phonological level, spreading will tend to turn an activation pattern in the phoneme nodes into a single morpheme, and particularly into morphemes that are semantically related to others in the sentence. As a result, sound errors tend to be words (or morphemes), which is the lexical bias effect, and these words (or morphemes) tend to be related to those in their vicinity, which is the semantic bias effect. These bias effects in phonological errors are discussed in detail in the section on phonological encoding.

The second general influence on errors is similarity. Similarity plays a role in speech errors in two ways. First, the interacting items in an error tend to be similar—interacting words tend to be similar in sound and meaning (Cutler, 1980; Dell & Reich, 1981; Fay & Cutler, 1977; Fromkin, 1971; Harley, 1984; Hotopf, 1980; Nooteboom, 1969), and interacting morphemes, phonemes, and other sound combinations tend to be similar in sound (e.g., Goldstein, 1977; Kupin, 1979; MacKay, 1970; Shattuck-Hufnagel & Klatt, 1979; Stemberger, 1982b). Second, the immediate environment of the interacting items tends to be identical. An example of environmental similarity is the repeated phoneme effect in sound errors (MacKay, 1970). Phoneme exchange errors are more likely when the sounds adjacent to the reversing ones are identical as in *left hemisphere* → *heft lemisphere*, in which the vowel /ɛ/ is adjacent to the reversing phonemes /l/ and /h/. In the theory, both item and environmental similarity can be understood as resulting from close connections in the lexical network. Items are similar if they connect directly to the same nodes at a lower or higher level. The morphemes *sink* and *some* are phonologically similar in that they both connect to the phoneme /s/; the words *swim* and *drown* are semantically similar because they (presumably) share conceptual nodes that capture their relation. Environmental similarity also is represented by common connections. The /I/ in *sink* and the /ʌ/ in *some* are environmentally similar in that they both connect indirectly to the node for /s/ via *sink* and *some*.

As was the case for output biases, similarity effects in errors also result from spreading activation. The closeness in the connection between similar items acts as a pathway for the activation of similar items to affect one another. An activated correct item will have similar competitors because of spreading, and if there are other factors that contribute to the activation of one of these competitors, the competitor's activation level may exceed that of the correct item leading to a slip. In general, spreading leads to correlations in the activation levels of nodes that are "near" to one another. Some of these correlations lead to output biases (the activation levels of a set of phonemes that compose a word tend to be correlated because of their connection to a common word or morpheme node), and some lead to similarity effects (items that share components will have correlated activation levels because of the common connections through the shared components).

The third general influence on error probability is the speaking rate. Both common sense and empirical data (Kupin, 1979; MacKay, 1971) tell us that errors are more likely when the speaker is speaking fast. This effect occurs in the theory because the rate of processing in the lexical network, reflected by spreading and decay rates, is assumed to be constant, regardless of the speaking rate. At fast rates, errors are common because there is just not enough time for the retrieval of the correct items. The old items are still activated because they have not had enough time to decay, and the new correct items have not received enough activation from the higher representations to compete successfully.

A final class of error effects concerns the distance over which items move in errors of misordering. In general, items are more likely to move short distances. Misordered sounds and morphemes tend to move to adjacent content words that are in the same phrase (Boomer & Laver, 1968; Garrett, 1975; MacKay, 1970). Misordered words move greater distances, possibly because the planning chunks at the syntactic level are larger, or because words can only move to appropriate syntactic slots (Garrett, 1975). The tendency for misordered items to move to adjacent or nearly adjacent positions is handled by the theory in a straightforward fashion. At any given time, the most highly active competitors for a given slot tend to be those that have just been, or are destined to be, selected in nearby slots.

This concludes the description of the general theory. Up to now, I have presented the theory's linguistic and processing assumptions and introduced in a general way how the theory accounts for various classes of error phenomena. In the remaining sections, the theory will be developed in detail—assumptions regarding the structure of the lexicon and generative rules are introduced and predictions are developed and tested. The next section deals with phonological encoding and the final section with morphological and syntactic encoding.

## Phonological Encoding

Phonological encoding can be defined as the processes by which the speech sounds that compose a morpheme or string of morphemes are retrieved, ordered, and organized for articulation. This section develops a model of phonological encoding that realizes the basic assumptions of the production theory presented earlier. The model accounts for existing findings, and generates and tests new predictions regarding sound errors. Because the model includes both the assumptions of the general theory and specific assumptions regarding the nature of phonology, some of the predictions are derived principally from the general theory, some from the hypothesized structure of phonology, and some from an interaction of the two. I will attempt to identify which is which.

### *Phonological Units*

Within the framework of the general theory, phonological encoding consists of the mapping between the morphological representation (a set of tagged morphemes) and a phonological representation, which up to now, has been pictured as a set of tagged phonemes. Why phonemes? Why not phonemic features or syllables, or for that matter, why not phonetic segments or phonetic features? There are two questions here: One concerns the size of the units of encoding, and the other, whether these units are essentially phonological or phonetic, that is, whether or not they represent nondistinctive allophonic variation.

It seems clear that sound errors occur at a level of representation that is phonological rather than phonetic. When sounds are misordered, they acquire the allophonic characteristics suitable for their new environments (Fromkin, 1971; Wells, 1951). For example, the /k/ in *Katz* is phonetically a front /k/, yet in the error *Fats and Kodor* for *Katz and Fodor,* the /k/ becomes a back /k/ accomodating to the back vowel /o/ in *Kodor.* This suggests that the unit that is misordered is phonological rather than phonetic and, more generally, that there exists a phonological representation in production and it is at this level that these kinds of sound errors occur. (For further evidence for the abstractness of the representation involved in sound errors see Crompton, 1981, and Stemberger, 1983.)

The question of the units of the phonological representation is also illuminated by speech-error data. Generally, there are two points of view. One can claim that the phoneme is *the* phonological unit, or one can claim that the phoneme is one of many units, including the syllable, syllabic constituents (consonant clusters, rime constituents), and features. The first claim is supported by the fact that the vast majority of sound errors involve the movement, deletion, substitution, or addition of a single phoneme. Nooteboom's (1969) analysis of sound errors, which is typical, found 89% single phoneme errors, 7% consonant clusters, and 4% for the remainder (other syllabic constituents and syllables). Although Nooteboom did not report any single feature errors, these have been found (Fromkin, 1971; Stemberger, 1982a). Because of the predominance of phoneme slips, some theorists characterize the phonological representation used in production as a set of phonemes (e.g., Fry, 1969; Shattuck–Hufnagel, 1979). Others (e.g., Fromkin, 1971) note the existence of sound errors involving other units and suggest that there is no single unit. A theory of phonological encoding has to address this issue that I call the *units problem.*

The units problem comprises two subproblems that I call the *syllable* and the *feature paradoxes.* The syllable paradox refers to the existence of conflicting evidence on the importance of the syllable as an encoding unit. On one hand is the rarity of syllable errors (Shattuck-Hufnagel, 1979), suggesting that the syllable does not have the status of a unit, if units are to be defined as speech-stream segments that frequently slip (e.g., exchange with others of their type). But on the other hand, the syllable seems to be very important because there is a syllable position constraint on sound errors. A moving sound nearly always retains its syllabic position when moving to another syllable.

The feature paradox is similar. First, there is the rarity of feature errors, which, like the case with syllables, signifies a very limited role for the feature as a unit (Shattuck-Hufnagel & Klatt, 1979). However, again like the syllable, features play an important role in determining which phonemes can slip with which. The interacting phonemes in exchange, anticipation, perseveration, and substitution errors tend to be similar in sound; that is, they share features. So, the features seem to be exerting their influence as units, but this is not revealed by the features themselves slipping, at least nowhere near to the extent of that of phonemes and syllabic constituents. Clearly, a model of phonological encoding must assign roles to the various linguistic units so as to account for these paradoxes.

### *A Simulation of Phonological Encoding*

This section presents a simulation model of phonological encoding. (The model is expressed as a simple Fortran program and is available on request.) Given an ordered set of morphemes it creates a second ordered set of phonological units. The model follows the outlines of the general theory introduced earlier and includes specific assumptions about phonological units. First, I will describe the relevant parts of the lexical network and the specific processing assumptions of the model. Then, I will present the results of simulation experiments in which the model encodes morphemes in its vocabulary. Its performance, in terms of the frequency and types of error, will be compared with known properties of sound errors.

*Lexical network.* The network proposed for phonological encoding consists of nodes for the following units: morphemes, syllables, syllabic constituents, (consonant clusters and rimes), phonemes, and features. Thus, the network contains a hierarchy of phonological nodes as exemplified in Figure 2. The nodes representing items connect to one another in a straightforward fashion, with two exceptions, the inclusion of special nodes for null elements and the syllabic position encoding for units that make up syllables.

Null elements stand for the absence of either an initial or a final consonant in a syllable. When a syllable has no initial consonant, it connects directly to an initial null element node as in the syllable /ən/, and when it has no final consonant, it connects to a final null element as in the syllable /spa/. These nodes are both convenient in that they simplify the phonological rules, as will be seen, and they constitute a hypothesis regarding the structure of syllables that may or may not be true. I will show later how certain error types can be interpreted as resulting from interactions involving null elements. However, it will also be seen that the null elements have shortcomings.

The second exceptional feature of the proposed network is the position encoding of consonant clusters, single consonants, and their features. Each of these node types is marked as to whether it is pre- or postvocalic (onset or coda). A unit that can occur in both positions must be represented by two nodes as can be seen with /k/ in Figure 2, one of which is marked as a potential onset and one as a potential coda. This position encoding of consonants is an ad hoc mechanism that enables the model to handle the positional specificity of interference, the fact that initial and final consonants rarely substitute for one another. Essentially, initial and final versions of the same consonant behave as if they are separate entities at this particular level in sentence production.

*Processing assumptions.* Processing in the network begins with the assignment of current node status to the morpheme tagged as first in the morphological representation. Its activa-

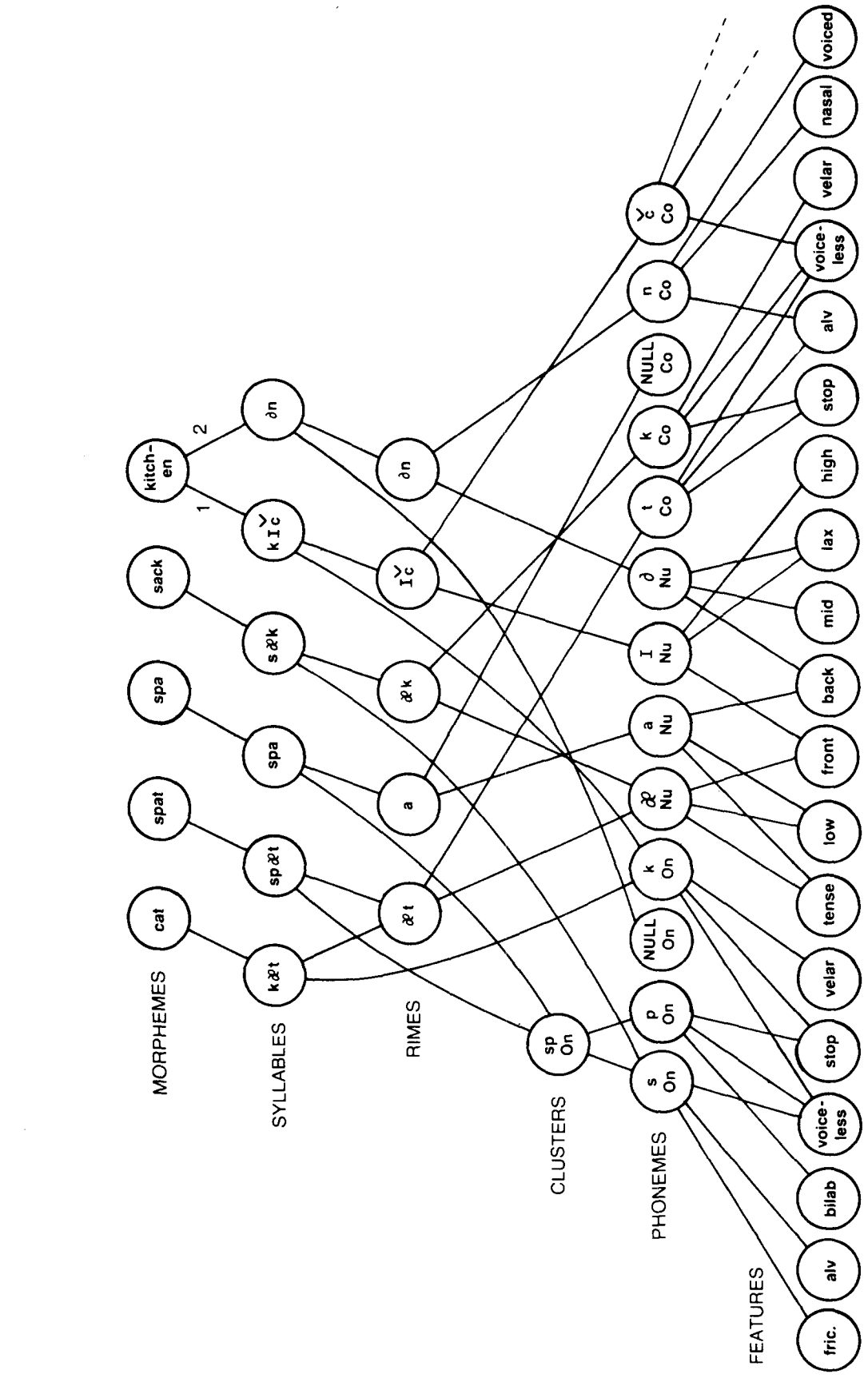


*Figure 2.* A piece of a network for phonological encoding in the theory. (Phoneme, cluster, and null element nodes are labeled as to whether they are potential (On)sets, (Nu)clei, or (Co)das. All connections are both top-down and bottom-up.)

tion level is then increased by 100 arbitrary units of activation (signaling activation). The activation level of the next morpheme in the representation is incremented by a lesser amount of activation called *anticipatory activation.* (This was always 50 arbitrary units.) The anticipatory activation is a special feature of the phonological encoding model that is included to represent the ongoing processing at the morphological and higher levels. Although these higher level encoding processes are not modeled here, their effects on the activation levels of the morpheme nodes can be crudely represented. One would expect that the upcoming morphemes in the same phrase as the current one would be activated due to the morphological and syntactic construction processes. The assumption that is adopted for the model, that the next morpheme after the current one receives activation, represents this expectation.

Following the adding-in of signaling and anticipatory activation, activation then spreads through the network according to the rule stated earlier. It is assumed that all downward connections (e.g., morpheme to syllable, syllable to phoneme) have a constant spreading rate, $p$, and that all upward connections have a constant value that is some fraction of $p$. After a fixed period of time, $r$ time steps, a single syllable is selected according to the following phonological rule:

$$\text{Syl} \rightarrow \text{Onset Nucleus Coda},$$

where *Onset* stands for either the initial null element, an initial consonant, or an initial consonant cluster; *Nucleus* stands for a vowel or diphthong; and *Coda* is either the final null element, a final consonant, or a final consonant cluster. Within each of the categories, the node with the highest activation level is selected and tagged as part of the phonological representation. When a node is tagged, its activation level is set to zero. (This is termed *postselection negative feedback.*) Parameter $r$, the number of time steps per syllable, represents the speaking rate. When $r$ is large, much time passes between the encoding of each syllable, and when it is small, little time passes. The time is taken up in the spreading of activation. To simplify matters, it is assumed that the onset, nucleus, and coda for a given syllable are selected simultaneously.

Before presenting further details of the simulation model, I will discuss the ramifications of the particular phonological selection rule previously mentioned. Clearly, it is not a complete description of English phonology. That is, it does not capture all of the regularities of the sound pattern. It can generate all possible English syllables, but it also generates many syllables that would be phonologically unacceptable in certain environments. For example, the rule does not take into account the word position of each syllable it selects. Certain syllables are unacceptable in word-final position such as those ending with the vowels /ɛ/ or /I/. In addition, the rule does not represent the dependencies between nucleus and coda in English, nor does it capture the phonological differences between stressed and unstressed syllables. Nonetheless, the rule chosen captures enough of the phonology to illustrate how the model works. Principally, the rule allows the model to select strings of sounds that do not appear in the lexicon, as well as the strings that do. In this way it both captures the productivity of the sound system and provides the opportunity for sound errors.

By adopting a selection rule whose categories are onset, nucleus, and coda, the model is taking a stand regarding the units of phonological encoding. Although the lexical network contains nodes standing for units of various sizes from morpheme to feature, only those units that can serve as onsets, nuclei, and codas are basic units. The phonological representation in the model consists of tagged nodes that stand for these three syllabic constituents. Most of the time these constituents are single phonemes, but for syllables other than CVC syllables, the tagged nodes include clusters and/or null elements.

A problem associated with a selection rule based on the syllable is that of multisyllabic morphemes. How many syllables are to be selected for each morpheme, and how are they ordered? For the simulation I assume that each syllable of a multisyllabic morpheme is marked for its morphemic serial position, as is the case for *kitchen* in Figure 2. When a given morpheme is the current node, one of its syllables will be the *current syllable.* Activation spreads between the current morpheme and its syllables in much the same way as that between other morphemes and their syllables, except that it spreads preferentially to the current syllable. The rule adopted for the model, which is designed to handle one or two syllable morphemes, is as follows: For one-syllable morphemes, activation spreads from the current morpheme to its syllable with spreading rate $p$, which is the same way as for the connection between a morpheme that is not current and its syllable(s). For two-syllable current morphemes, activation spreads from the morphemic node to its current syllable with the rate of 1.5 $p$, and to its other syllable with a 0.5 $p$ rate. So, the total activation sent in each time step from the current morpheme to its syllables is 2 $p$, which is also the total amount sent from a noncurrent two-syllable morpheme to its syllables ($p$ to each syllable). The only difference is that the activation flows preferentially to one syllable of the current morpheme. Thus, the order of the syllables, like that of the morphemes, is transferred from the morphological to the phonological representation through variations in activation level. As is the case with the current morpheme, the current syllable must be changed when appropriate. As each syllable is selected, either the current syllable changes to the next syllable of the current morpheme or, if the last syllable of the morpheme had been current, the current morpheme status changes to the next morpheme. A serial order mechanism that uses inhibitory links between the syllables such as that proposed by MacKay (1982) might deal with this issue in a simpler manner. However, because the issue of syllabic order within a single morpheme is not directly concerned with the model's empirical domain, I set the issue aside.

*Simulation network and parameters.* Fifty single-morpheme nouns were selected at random from the Toronto word pool, a collection of 1,024 two-syllable common English words not more than eight letters long, to provide the simulation's vocabulary. The resulting network of morphemes, syllables, syllabic constituents, phonemes, and features contained 251 nodes with 941 connections among them. The parameter values for the simulation runs were (a) signaling activation, 100 units of activation added to the current morpheme when it first obtains current status; (b) anticipatory activation, 50 units added to the next morpheme; (c) speaking rate, $r$ time steps per syllable, varies in integral units from 2 to 8; (d) spreading rate, $p$, 0.3 per time step; (e) decay rate, $q$, 0.6 per time step; and (f) spreading rate for upward connections, 0.5 $p$. These values were, by and

Table 3
*Simulated Effect of Speech Rate and Number of Words in Utterance on Error Probability*

| No. of words in utterance | No. of time steps per syllable ($r$) | | | |
|---|---|---|---|---|
| | 2 (fast) | 3 | 4 | 5 (slow) |
| 1 | 61.2 | 0.0 | 0.0 | 0.0 |
| 2 | 61.9 | 11.3 | 1.3 | 0.5 |
| 6 | 60.0 | 24.3 | 5.7 | 0.6 |

*Note.* Numbers are the percentages of phonemes that were not correctly encoded.



large, not set in a principled way. The exceptions were the spreading and decay rates, which were selected so that, given the other parameters, activation did not grow without bound and the simulation could correctly encode each morpheme in its vocabulary when spoken singly with a rate of three time steps per syllable.

*Effect of speaking rate and utterance length on error rate.* The first simulation experiment varied the number of words encoded and the speaking rate. For the simulation to behave like a human speaker, it should err more often when speaking quickly and when it has too many words to say. The simulation encoded strings of one, two, or six words (randomly selected with replacement from its vocabulary) with $r$ at 2, 3, 4, or 5. There were 216 words attempted in each condition, and percentages of incorrect phonemes are given in Table 3. First consider performance when $r = 2$. Here the simulation does very poorly because the vowel and final consonant nodes are three network links away from the morpheme nodes, and two time steps is not enough for an activated morpheme to influence these phonemes. With $r > 2$, there is enough time for single words to be encoded without error (recall that parameters were set so that this would be true). Longer word strings, however, had occasional slips—the longer the string, the greater the probability that a given sound would slip. The effect of speaking rate was also as it should be. As the simulation's speech slowed, errors became less likely for reasons discussed earlier.

The next simulation experiment will be used to reveal the types of errors made by the model and their frequencies.

*Error types in the simulation.* Unlike the previous simulation runs, these runs included random fluctuations in activation levels. The activation level of every node was changed every time step by adding in a normal deviate of $M = 0$ and $SD = .2A$ ($j$, $t$). Thus, the standard deviation of the noise added to each node $j$ was proportional to that node's activation level at that time. The noise serves to model the many influences on activation level that come from other nodes not represented in the simulation's network. Earlier, I referred to this as linguistic background noise.

The simulation encoded 120 word pairs randomly selected from its vocabulary at three speaking rates, $r = 3$, 4, and 8. Errors were identified and coded by type of error (exchange, anticipation, etc.) and by size of the slipping unit (phoneme, consonant cluster, etc.), as is usually done with error data (e.g., Fromkin, 1971; Garrett, 1975). If a given error could be classified in more than one way, the simplest classification was chosen. For example, the target *menu sucker* (/menyuw sʌkər/) being encoded as /menyər sʌkər/ could be categorized as the anticipation of the vowel /ə/ replacing the vowel /uw/, together with the anticipatory addition of /r/. However, the simplest categorization (that requiring the smallest number of basic errors) is that the rime constituent /ər/ was anticipated replacing the rime constituent /uw/. Another possible ambiguity in coding concerned anticipations and perseverations. Errors such as *infant urchin → infin urchin* could be classified as either the anticipation or the perseveration of the /In/ constituent. From the point of view of the simulation, these errors are caused by both anticipatory and perseveratory sources of activation and, hence, they are coded in a separate category called *anticipation/perseveration.*

In general, the simulation's errors exhibited many of the properties present in collections of slips of the tongue. Table 4 gives some of the errors, and Table 5 presents the frequencies of the error types at the different speaking rates. To a first approximation, the obtained errors look like slips of the tongue. When sounds move from one syllable to another, they retain their syllable position because of the position encoding of sounds in the network. In addition, the resulting syllables are phonologically acceptable. That is, they conform to the selection rule that governs the construction of the phonological representation. The errors deviate in their appearance from normal speech errors in two significant respects. The simulation's errors are characterized by movement of sounds between syllables regardless of the syllables' stress values. True errors tend to occur in stressed syllables, and moving sounds tend to go to syllables that have the same stress value as the syllable from which the sounds originated. Because stress was not represented, the effect of stress on errors was not present. A second deviant aspect of the errors is that they would occasionally create syllables that were unacceptable in word-final position (e.g., *dowry* → /dawrI/). This happens because the simulation does not "know" the word position of the syllables that it selects. It just selects syllables, one after another. These unusual features of the errors' appearance point out the incompleteness of the model's representation. Despite this, the simulation is able to model some aspects of the types of errors that occur and their frequencies.

With the exception of unambiguous feature errors, noncontextual errors, and errors involving the deletion of a syllable, the

Table 4
*Examples of Errors From Phonological Encoding Simulation*

| Error | Type |
|---|---|
| *Kinsman → minsman* | Phoneme anticipation |
| *Infant → anfint* | Phoneme exchange |
| *Ensign mercy → mensign mercy* | Anticipatory addition |
| *Crackle → cackrel* | Phoneme shift (note division of consonant cluster/kr/) |
| *Topsail → topsay* | Phoneme deletion |
| *Topsail → topsop* | Perseveration of rime constituent |
| *Typist figure → typig fisture* | Exchange of cluster and single phoneme |
| *Seaman woman → seaman moman* | Phoneme anticipation/ perseveration |
| *Content convict → con con convict* | Syllable anticipation/ perseveration |
| *Topsail beggar → boptail beggar* | Multiple-site error |

Table 5
*Frequencies of Error Types From Simulation at Different Speech Rates*

| | Error type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Unit size | A | P | AP | E | D | AA | PA | S | Total |
| | | | | $r = 3$ (time steps per syllable) | | | | | |
| Phoneme | 29 | 53 | 12 | 8 | 36 | 5 | 6 | 4 | 153 |
| Cluster | 1 | 1 | 1 | 1 | 8 | 0 | 0 | 2 | 14 |
| Rime | 6 | 50 | 0 | 2 | —[a] | — | — | — | 58 |
| CV | 2 | 0 | 0 | 0 | — | — | — | — | 2 |
| Syllable | 0 | 3 | 1 | 0 | — | — | — | — | 4 |
| Total | 38 | 107 | 14 | 11 | 44 | 5 | 6 | 6 | 231 |
| | | | | $r = 4$ | | | | | |
| Phoneme | 29 | 14 | 8 | 2 | 14 | 4 | 1 | 0 | 72 |
| Cluster | 1 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 5 |
| Rime | 6 | 1 | 4 | 0 | — | — | — | — | 11 |
| CV | 1 | 0 | 0 | 0 | — | — | — | — | 1 |
| Syllable | 0 | 0 | 0 | 0 | — | — | — | — | 0 |
| Total | 37 | 16 | 12 | 2 | 17 | 4 | 1 | 0 | 89 |
| | | | | $r = 8$ | | | | | |
| Phoneme | 46 | 1 | 8 | 0 | 17 | 6 | 1 | 0 | 79 |
| Cluster | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| Rime | 4 | 0 | 0 | 0 | — | — | — | — | 4 |
| CV | 2 | 0 | 0 | 0 | — | — | — | — | 2 |
| Syllable | 1 | 0 | 0 | 0 | — | — | — | — | 1 |
| Total | 55 | 1 | 8 | 0 | 17 | 6 | 1 | 0 | 88 |

*Note.* A = anticipation; P = perseveration; AP = anticipation/perseveration; E = exchange; D = deletion; AA = anticipatory addition; PA = perseveratory addition; S = shift; CV = consonant–vowel.
[a] Cells with a bar are necessarily zero because the model cannot delete or add a vowel. Parameter values are given in the text.

model generated the basic types of sound errors as can be seen from Table 5. An unambiguous feature error is an error in which a single feature moves, or exchanges places with another, and the error cannot be classified as a phoneme error. For example, Fromkin (1971) reported the errors *glear plue sky* for *clear blue sky.* This is a true feature error—the unvoiced feature of the /k/ in *clear* and the voice feature of the /b/ in *blue* have exchanged places, creating /g/ and /p/. Because there are no other /g/s or /p/s in the environment of the slipping sounds, it seems correct to characterize the error as the exchange of features. Errors involving phonemes that differ by a single feature such as *gall the curl* for *call the girl* could be feature errors (exchange of the voicing and unvoicing features), but these are usually classified as phoneme errors (exchange of /k/ and /g/). The simulation made errors of the latter type but not the former. It is possible for an unambiguous feature error to occur, but it would be unlikely in the model because selection takes place at the level of the phoneme and consonant cluster.

The fact that the simulation made no noncontextual errors is not a problem for the model. Increasing the linguistic background noise would result in some errors whose source was outside of the intended utterance. The absence of syllable deletions in the simulation's output points to the fact that the model cannot drop a syllable. It posits a very simple relation between the morphological representation and phonological frame building—the phonology selects a syllable for every syllable present in the set of planned morphemes.

The errors that the simulation made can be placed in two categories, those in which the error does not change the number of segments (substitutions) and those in which the error does (nonsubstitutions). The former category includes anticipations, perseverations, and exchanges. The latter category is made up of deletions, anticipatory and perseveratory additions, and shifts. This categorization, however, obscures the fact that substitutions and nonsubstitutions are essentially the same thing in this simulation model. A nonsubstitution error is simply a substitution that involves either the null element and a consonant, or a single consonant and a cluster including that consonant. For example, a shift is either the exchange of a consonant and a null element (*urchin marriage* → *murchin arriage*) or the exchange of a consonant and a cluster containing it (*infant* → *intfan,* exchange of /nt/ and /t/). Anticipatory and perseveratory additions are analogous to anticipations and perservations in the same way. Deletions occur when a null element substitutes for a consonant or a consonant substitutes for a cluster that contains it. Thus, the chosen selection rule, which selects for null elements and consonant clusters, enables the model to produce both substitutions and other errors.

To conclude the discussion of the form of the model's errors, I should note that occasionally the model produced multiple-site errors such as *topsail beggar* → *boptail beggar* that seem to be related to exchanges. According to the error coding scheme I used, this would be listed as two errors, the anticipation of /b/ and the perseveration of /t/. However, this misses the subtle causal relations between the errors. The anticipation of /b/ sets the stage for the perseveration of /t/ because the /t/, which was not selected and turned off by postselection negative feedback, remains active and thus can replace the onset /s/ in the next syllable. Notice that this is similar to what happens in an exchange such as *topsail* → *soptail.* The difference is that in an exchange the replaced sound /t/ perseverates to the location of the replacing sound /s/. In the model, exchanges are more likely than other multiple-site errors because, in an exchange, the application of postselection negative feedback to the replacing sound /s/ creates a weakened /s/ and so the replaced sound /t/ has a natural place to go. The /s/ was not weakened by this feedback in the error *topsail* → *boptail.* Evidence that errors do involve these kinds of causal interactions is provided by the fact that error collections contain a few examples of multiple-site errors that are not exchanges, such as *motor cycle ride* → *rotor mycle ride* (from Shattuck-Hufnagel, 1979).

Other than producing some types of sound errors, the model produces many of the relations among the frequencies of the types. Table 6 compares the simulation's errors with those of Nooteboom's (1969) corpus of Dutch errors and Shattuck-Hufnagel's (1983) English exchange errors. The agreement between the model's error percentages and those of the corpora is very good when the model's speaking rate is four or eight. Some of this similarity can be attributed to serendipitious factors, such as judicious selection of parameter values. However, there are a number of effects that come directly from the model's structure.

The model closely mimics the relative frequencies of the sizes of the slipping units. Single-phoneme slips predominate in both the actual and simulated data. In the simulation the predomi-

Table 6
*Comparison of Simulation Errors With Sound-Error Collections*

| Speech rate | Collection: Simulation $r = 3$ | Simulation $r = 4$ | Simulation $r = 8$ | Nooteboom (1969, Dutch) | Shattuck-Hufnagel (1983, English)[a] |
|---|---|---|---|---|---|
| | | | Unit size | | |
| Phoneme | 66 | 81 | 90 | 89 | 70[b] |
| Cluster | 6 | 6 | 2 | 7 | 16 |
| Vowel–consonant | 20 | 10 | 5 | 2 | 6 |
| Consonant–vowel | 1 | 1 | 2 | 0.5 | 2 |
| Feature | 0 | 0 | 0 | 0 | 1 |
| Syllable | 2 | 0 | 1 | 1.5 | 3 |
| Other | 5 | 2 | 0 | | 2 |
| Total | 100 | 100 | 100 | 100 | 100 |
| | | | Error type | | |
| Anticipation[c] | 19 | 48 | 67 | 57 | |
| Perseveration[c] | 49 | 25 | 6 | 13 | |
| Exchange | 5 | 2 | 0 | 5 | |
| Total substitution | 73 | 75 | 73 | 75 | |
| Deletion | 19 | 19 | 19 | 3 | |
| Addition | 5 | 6 | 8 | 21 | |
| Shift | 3 | 0 | 0 | 1 | |
| Total nonsubstitution | 27 | 25 | 27 | 25 | |

*Note.* All numbers are percentages of the total number of errors in the collection. The simulation data come from the same runs as in Table 5.
[a] Exchanges only.
[b] This percentage includes Shattuck-Hufnagel's (1983) VV category.
[c] The anticipation/perseveration errors from the simulation have been alloted one half to the anticipation category and one half to the perseveration category.



nance of phoneme slips reflects the fact that selection takes place by onsets, nuclei, and codas, which tend to be individual phonemes. However, groups of phonemes tend to be selected together if there is a higher order node connected to them. Thus, in the model, because vowel and final consonants connect directly to the higher order node that represents the rime constituent, they tend to slip together. This is seen in the greater number of VC as opposed to CV slips in the model's data. The fact that this is present in the actual data supports the model's treatment of these constituents. In general, higher order constituent nodes lead to a positive correlation in the activation levels of the nodes connected to them. This results in the tendency for the phonemic and cluster nodes that form a single constituent to be either selected together or not selected together. This tendency, however, is stronger with rimes than with syllables. Although there is a node for each syllable, this node is two network links away from its vowel and final consonant nodes. As a result the higher order constituent effect is limited in the case of syllables. This is seen in the simulation's output, which has few whole syllable errors.

Several investigators have pointed out that the relative proportion of the sizes of the slipping units in sound errors points to a hierarchial syllable structure with the main break between onset and rime and a secondary break between nucleus and coda (e.g., MacKay, 1972; Shattuck-Hufnagel, 1983; Stemberger, 1983). Other evidence that does not involve speech errors supports this position, as well (e.g., Treiman, 1983, 1984). The model presented here provides the first formal demonstration that the inclusion of units representing syllabic constituents leads to the correct proportion of errors in terms of unit sizes. The aspects of the model that make this possible are, first, the tendency for positively correlated activation levels for closely connected units (which comes directly from the general theory of production), and second, from the assumption that selection is guided by rules whose categories are onset, nucleus, and coda.

As mentioned before, the simulation made no unambiguous feature errors. Yet the inclusion of feature nodes in the network does influence the errors. In particular, it creates a tendency for similar phonemes to slip with one another. Two phonemes sharing one or more features tend to have positively correlated activation levels. When a competing phoneme is similar to the correct one, it gains activation because of its similarity to the correct one, and thus has an increased chance of replacing it. This effect can be seen indirectly in the simulation's output. When the speaking rate is three time steps per syllable, there is not enough time for the feature nodes to influence activation levels. When this is the case the mean number of features shared between the interacting phonemes is 0.7. Decreasing the speaking rate (and, hence, increasing the amount of spreading activation) increases the similarity of the interacting phonemes in the slips. With $r$ at 4 and 8, the mean number of shared features

was 0.8 and 1.1, respectively. Each phoneme was characterized by three features in the simulation, so the similarity effect was not large. The effect can be increased, however, by increasing the spreading rates for the upward feature-to-phoneme connections. Thus, the model can account for the phonemic similarity effect. It does so because of the inclusion of feature nodes and through the basic spreading activation assumptions.

So far, it has been shown that the model mimics the frequencies of slips with respect to the size of the slipping units. In so doing, the model offers a solution to the units problem and its accompanying syllable and feature paradoxes. Recall that the syllable paradox refers to the rarity of syllable errors, despite the role of syllable position in governing phoneme errors. In the model, the role of the syllable is primarily that of a schema of selection (the rule *Syl* → *Onset Nucleus Coda*) rather than a unit of selection (see Shattuck-Hufnagel & Klatt, 1979). For an entire syllable to slip, the onset, nucleus, and coda must all be wrongly selected, which is unlikely. The model's treatment of features resolves the feature paradox as well. Features are present in the lexical network, thus leading to the tendency for similar sounds to slip. However, selection does not take place at the feature level. As a result, an error that looks like an unambiguous feature error would be rare.

Now I will discuss the proportions of the types of errors in the simulation, independently of the size of the slipping units. As Table 6 shows, these proportions are in fairly close agreement with those from Nooteboom's (1969) error corpus. However, the proportions of the types in the model can be pushed around to some extent by varying the model's parameters. For example, slow decay favors exchanges and perseverations at the expense of anticipations. Also, the chance of an exchange or anticipation can be increased by increasing the amount of anticipatory activation or extending it more than one word in advance of the current one. Nevertheless, in the model there are quantitative relations among the error types that are independent of parameter values and therefore provide a test for the model. For example, exchanges are necessarily less common than anticipations or perseverations in the model. This appears to be true in error corpora as well (e.g., Garnham, Shillcock, Brown, Mill, & Cutler, 1981; Nooteboom, 1969; Stemberger, 1982a). A basic requirement for an anticipatory slip to become an exchange is that there must be a significant tendency for replaced sounds to stay activated. Creating this tendency in the model (by decreasing the decay rate) also produces a number of simple perseverations. However, making this perseverative tendency very strong, so that most anticipations become exchanges, causes the model to make huge chains of perseverations resulting in complete gibberish. So, in general, setting parameters so that the model is reasonably fluent leads to anticipations and perseverations being more common than exchanges. Note that the same is true for anticipatory additions, perseveratory additions, and shifts, which are (in the model) special types of anticipations, perseverations, and exchanges, respectively.

Another quantitative effect that is independent of parameter values is the relative number of substitution versus nonsubstitution errors. In the model, a nonsubstitution (addition, deletion, shift) is simply a substitution involving certain types of onsets and codas. Thus, the relative proportion of nonsubstitution errors is determined by the structure of the syllables in the simulation's vocabulary. The simulation predicts that nonsubstitution errors should be less frequent than substitutions simply because English (and Dutch) syllables frequently have single-consonant onsets and codas. So, when one onset, nucleus, or coda replaces another it is usually a single phoneme replacing another single phoneme.

Finally, I should point out that the simulation differs substantially from the error corpus with respect to deletions and additions. The simulation tends to delete—unlike Nooteboom's (1969) and other corpora (e.g., Stemberger, 1984c), in which additions are more common. The model tends to delete because the null elements are very frequent "sounds"; that is, there are many connections to them. When activated, they send activation to many other nodes, which, in turn, send activation back again. Hence, over time they become more highly activated than do their less frequent competitors. The result is an asymmetry—the null element substitutes for a consonant creating a deletion, more often than a consonant substitutes for the null element creating an addition. I return to this point later, when problems in the model are discussed.

### *Spreading Activation and Sound Errors*

So far, it has been demonstrated by simulation that the phonological encoding model produces some of the types of sound errors that occur in error collections, mimics some facts about the proportions of these types, and has offered resolutions to the syllable and feature paradoxes. The theoretical work has been done through a combination of the general theory's assumptions with some specific assumptions about the representation and ordering of speech sounds.

This section explores effects on sound errors that can be explained primarily through the theory's general assumptions about spreading activation. After discussing these effects, which include output biases, repeated-phoneme effects, and speaking-rate effects, I will present the results of a comprehensive experimental test of the model's predictions—a test that simultaneously varies several factors that are known to affect error probability. Then a new simulation model, particularly designed to mimic the experimental test, will be used to fit the data.

*Output biases.* Because the model's activation spreads both downward (from morphemes to phonemes and features) and upward (from phonemes and features to morphemes), a complex pattern of positive feedback loops is created as each syllable is encoded. Generally speaking, this causes the activation levels to adjust themselves so as to reflect the entire stored vocabulary in interesting ways. A number of specific effects occur.

First, the activation pattern among the sound nodes adjusts itself over time so that the most highly activated nodes correspond to a single morpheme or word. The resulting effect on the model's errors is lexical bias, the tendency for errors to create existing morphemes or words. More so than any other error phenomenon, lexical bias directly reveals the interactive nature of the model. At the sound level of processing, adjustments that reflect higher level knowledge are made through spreading activation. The model acts as if there is a built-in editor that prevents nonword strings from being encoded. The lexical-bias effect has previously been interpreted as evidence that the speech plan is edited before being articulated. Baars et al.

(1975) argued that the tendency for slips to make words shows that a lexical editor scans potential output and vetoes nonword outcomes. Spreading activation, however, accomplishes this function automatically without any special mechanism.

In addition to explaining lexical bias, the model explains or predicts many related output biases in much the same way. These are briefly listed as follows:

1. Syllable bias. Because each existing syllable has a single node in the network, there will be a bias for errors to create existing syllables (e.g., /spæč/, as in *spatula*) over phonologically legal but nonoccurring syllables (e.g., /splč/).

2. Frequency biases. A sound that is present in many different syllables, such as initial /t/, has many more connections from its node to syllable nodes than one that occurs in few syllables, such as initial /f/. The number of these connections to a sound (its syllabic frequency) has consequences for its activation level. As activation spreads upward from sounds to syllables and back again, frequent sounds get more activation from activated syllables than do infrequent sounds. This creates a bias such that frequent sounds tend to replace infrequent ones more than the other way around.

There should be a similar frequency bias toward syllables that occur in many morphemes, and morphemes that occur in more than one word. Sound errors should, in general, create units whose nodes have many connections to them.

3. Contingent frequency bias. Just as spreading activation favors the utterance of units that are frequent, it favors the production of combinations of units that are frequent. For example, in the context of an initial /k/, the vowel /æ/ is more frequent than /U/; that is, there are more syllables that begin /kæ/ than /kU/. Given an initial /w/, however, the contingent frequencies of /æ/ and /U/ are not very different from each other. In the model, a highly activated /k/ node would bias for the high activation of /æ/ relative to /U/ because of feedback from the many /kæ/ syllables. If /w/ is the dominant initial consonant, there would not be this bias because the number of /wæ/ and /wU/ syllables is not so different. Slips should thus reflect these biases by creating high-frequency sound combinations.

The output biases that the model predicts have not been explored empirically to a great extent. As I mentioned earlier, the lexical-bias effect has been demonstrated both experimentally (Baars et al., 1975) and in a collection of natural errors (Dell & Reich, 1981), although its status is controversial in natural errors (e.g., Garrett, 1976). The predicted bias for frequent sounds replacing infrequent ones has been found in experimentally elicited sound errors (Levitt & Healy, 1985; Motley & Baars, 1975, Experiment 2). Motley and Baars (1975, Experiment 3) also found that initial consonants with a high contingent frequency (contingent on the following vowel) tend to replace those with a low contingent frequency, even if the absolute frequencies of the consonants are the same. Thus, there is support for the contingent frequency prediction of the model, as well.

Motley and Baars (1975) interpreted their finding of output bias as support for a pre-articulatory editor that is sensitive to frequency and contingent frequency. The phonological encoding model presented here explains these biases differently. As was the case with lexical bias, spreading activation produces the effects automatically. In general, the model finds it easiest to encode strings of sounds that are likely to be correct—words, morphemes, and syllables that exist in its vocabulary, and frequent sounds and sound combinations.

*Repeated-phoneme effect.* In addition to accounting for output biases, the model also accounts for some similarity effects in errors. Earlier it was shown using the simulation that similar sounds (phonemes sharing features) tend to slip with each other in the model. Another similarity effect in sound errors is the repeated-phoneme effect—repeated sounds tend to induce the misordering of the sounds around them (Dell, 1984; MacKay, 1970; Nooteboom, 1969; Wickelgren, 1969). For example, the repeated /æ/ in the phrase *hat pad* could induce either an exchange *pat had,* an anticipation *pat pad,* or a perseveration *hat had.* Existing theoretical accounts of this effect attribute it to contextual influences between adjacent phonemes; that is, both /h/ and /p/ are fitted to go with a following /æ/, thus making their rearrangement more likely (MacKay, 1970; Wickelgren, 1969).

In recent research the contextual influences explanation of the repeated-phoneme effect has been questioned. It has been found that repeated phonemes can induce the misordering of sounds that are not adjacent to the repeated ones, as in *hat pet* → *pat het,* in which the repeated final /t/ induces an initial consonant slip. This generalization of the effect to nonadjacent sounds has been verified through experimentally elicited errors as well as observed in an error corpus (Dell, 1984).

The spreading-activation model explains the repeated-phoneme effect as a consequence of the structure of the network. Two syllable nodes that share a sound connect indirectly through the node for the common sound. This leads to a positive correlation in the activation levels of nodes for the other sounds in both syllables. Thus, it becomes more difficult to keep sounds from the wrong syllable from getting a high activation level and being selected, thereby creating a misordering error of some kind. The repeated sounds can affect any of the other sounds in the syllables. They need not be directly adjacent to the sounds that they move, because the effect is mediated by higher order syllable nodes rather than by phoneme-to-phoneme connections. In this way the model directly explains both the general effect of repeated sounds and the fact that it is not limited to phonemes right next to the repeated ones.

*Speaking-rate effects.* Earlier I showed that the simulation's error rate depends on the speaking rate (number of time steps per syllable), such that there is a simple trade-off between speed and accuracy of encoding. The reason for this is that the parameters of activation dynamics (spreading and decay rates) are constant, regardless of the speaking rate. At slow rates there is plenty of time both for activation from the current morpheme to spread to the correct sounds and for activation of preceding sounds to decay. Reducing the available time per syllable by increasing the rate makes less time available for constructive spreading and decay, which results in the wrong sounds being more highly activated than the correct ones.

More interesting than the model's account of the general speed–accuracy trade-off are its specific predictions regarding error patterns at different rates. These predictions are perhaps the most revealing of the model's structure because they tie effects due to spreading, such as output bisaes and similarity effects, to the time that is available for spreading, which is governed by the speaking rate. I will discuss three predictions of

this type—interactions with speaking rate involving lexical bias, the repeated-phoneme effect, and different types of error.

1. Lexical bias and speaking rate. Because spreading is the mechanism for output biases, the more spreading that can occur, the stronger the bias should be. Thus, when the speaking rate is slow there should be a greater tendency for errors to create existing morphemes, syllables, and frequent sounds and sound combinations. I will limit the discussion to lexical bias, but it will apply for other output biases as well.

Consider a situation in which the morpheme *kitten* is the current morpheme, and for some reason the phoneme node for initial /d/ is highly activated. As activation spreads from *kitten* downward (preferentially to the first syllable *kit*), the sounds initial /k/, /I/, and final /t/ are highly activated (being connected to the first syllable), and the nodes for the constituents of the second syllable have a lesser degree of activation.

Let us assume that early in this process the activation of the wrong sound, initial /d/, is higher than the correct one, initial /k/. If selection occurs at this point, the slip *kitten* → *ditten* happens. (Whether this is an anticipation, perseveration, part of an exchange, or a noncontextual substitution error depends on where the /d/ got its activation.) Had activation been allowed to spread longer (that is, if the speaking rate were slower), the activation of /d/ would have decreased relative to /k/ because there is no morpheme node for *ditten* to support its continued activation. The pattern of activation that favors the selection of *ditten* is not stable in the network because the feedback from the sound nodes converges on the morphemes *kitten, mitten,* and so forth, which will, in turn, activate initial /k/ and /m/ at the expense of initial /d/. Thus, if activation is given time to spread (slow speaking rate) lexical bias will come into effect and edit out the sound or sounds that are not compatible with a single morpheme. Had the contaminating sound been /m/, not /d/, the /m/ would not have been edited out because the morpheme *mitten* would have supported its activation. Thus, at a slow speaking rate the *kitten* → *mitten* slip would be more likely than the *kitten* → *ditten* slip.

2. The repeated-phoneme effect and speaking rate. Like the lexical-bias effect, the repeated-phoneme effect is brought about by activation spreading both upward and downward in the lexical network. Activation from the current syllable spreads to another syllable if it shares a sound with the current syllable via the node for the shared sound. This increases the chance of a slip involving those syllables. At slow speaking rates, there is simply more time for this interaction to occur. So, the repeated-phoneme effect should be stronger at slower speaking rates.

3. Error types and speaking rate. The model's exchange errors (e.g., *pancake* → *canpake*) happen when the replaced sound of an anticipatory slip, such as /p/ in the example, remains active enough to replace the correct sound in the next syllable. If the replaced sound's activation decays, and thus it is not selected, the slip is an anticipation, *pancake* → *cancake.* Because decay is a function of time, the relative proportion of exchanges and anticipations should be affected by the speaking rate. In particular, the ratio of exchanges to anticipations should decrease as speech slows.

So much for the model's account of output biases, repeated-phoneme effects, and speech-rate effects. By presenting these effects in some detail, I hope to give a feel for the variety of phenomena that characterize the model's errors. Some of these phenomena have support from experimental studies of errors. These include lexical bias, sound frequency and contingent frequency biases, the repeated-phoneme effect for both adjacent and nonadjacent sounds, and the main effect of speech rate on errors. The fact that the model creates these effects directly through its spreading-activation assumptions is strong support for these assumptions. Other error phenomena created in the model have not been investigated empirically and thus are predictions from the model and the general theory behind it. These predictions center on interactions between the effects mentioned above and the speaking rate. Simply put, the pattern of errors should differ at fast and slow rates.

The experiment reported in the next section will test these predictions by looking at the lexical-bias and repeated-phoneme effects at three different speaking rates for three different error types: initial-consonant exchanges, anticipations, and perseverations.

Table 7
*Slip Procedure for Error Generation*

| Type of pair | Subject sees | Cued? | Subject is supposed to say | Subject says |
|---|---|---|---|---|
| Interference | Bid meek | No | Nothing | Nothing |
| | Bud muck | No | Nothing | Nothing |
| | Big men | No | Nothing | Nothing |
| Critical | Mad back | Yes | "Mad back" | "Mad back," "bad mack," "bad back," or "mad mack" |

## *Experimental Test of the Model*

*Error-induction technique.* This experiment generated initial-consonant misordering errors under controlled conditions. The technique for creating errors was a modified version of the Slip procedure, originally developed by Baars and Motley (1974), and has since been used in many experiments (e.g., Baars, Motley, & MacKay, 1975; Dell, 1984; Motley & Baars, 1976; Stemberger & Treiman, 1986). In the procedure used by Baars and Motley, subjects are presented visually with word pairs at a 1-s rate. Immediately after the presentation of some pairs—the cued pairs—a signal of some sort is given that directs the subject to say the most recently presented word pair aloud as quickly as he or she can. Certain word pairs, the *critical cued pairs,* are each preceded by from three to five *interference pairs* that are not cued, but serve to bias the subject to make a slip when saying the critical cued pairs. See Table 7 for an example.

Notice in Table 7 that the interference pairs set up the subject to expect a word beginning with /b/ followed by a word beginning with /m/. The effect is to induce an initial-consonant exchange, anticipation, or perseveration on about 10% of the trials. So the procedure is quite effective at producing the desired behavior, namely slips. But are these truly speech errors, that is, errors of output? One might object to the technique on the ground that the errors are not real slips but instead are due to subjects either misreading the words or failing to remember them. Baars and Motley (1974) reviewed the evidence and con-

cluded that the errors are truly slips. I will not evaluate their arguments. Instead I have introduced changes in the procedure to weed out errors that are not slips.

In the present experiment, after a subject said a cued pair, he or she had to state whether they had made an error, and had to repeat slowly the word pair that they should have said. If a subject detected the fact that they had erred, and correctly repeated the intended word pair, then whatever the subject spoke was classified as a genuine speech error.

A second change in the original procedure was the inclusion of a deadline. This was how the speaking rate was manipulated. After the critical pair was presented, a signal for the subject to speak occurred. (This was a series of question marks.) Then, after either 500 ms, 700 ms, or 1,000 ms, a tone sounded. Subjects were instructed to say the word pair before the tone. The deadline interval (the time between the question marks and the tone) was held constant for a given subject. The subjects were given practice with their deadline so that they knew how much time they had to speak.

Subjects who were assigned a short deadline had to speak quickly. They had to both begin the word pair quickly and articulate rapidly in order to finish before the deadline. Pilot work showed that a 500-ms deadline was as fast as most subjects could go and still speak coherently. If the deadline was longer, subjects took longer to begin speaking and spaced out the words to a greater extent.

*Experimental factors.* The principal goal of this experiment was to obtain a large number of errors under many conditions in order to give the model a rigorous test. The experiment combined factorially three variables: (a) whether the error outcome was a word, (b) the presence or absence of a repeated phoneme (the vowel) in the pair of stimulus words, and (c) the speaking rate, as determined by different deadlines (500 ms, 700 ms, and 1,000 ms). Some predictions from the model are as follows: (a) Errors creating words should be more likely than those creating nonwords, particularly at longer deadlines. (b) Word pairs with a repeated phoneme should lead to more slips than should those with no repeated phoneme, but this effect should also be more in evidence at longer deadlines. (c) The error rate as a whole should be lower at longer deadlines, and this should be particularly true for exchanges, which tend to become anticipations as speech slows.

Overall, the experiment was designed to generate 36 error proportions (12 conditions × 3 error types) in order to provide a comprehensive set of data for use in model fitting.

*Materials.* As dictated by the Slip procedure, the stimuli were word pairs. Each word was a single syllable, and 98% of the critical cued pairs were CVCs with 2% CVCCs. The critical cued pairs were constructed as follows. There were 20 sets of four pairs each. The four pairs within a set were designed to manipulate two of the experiment's independent variables—the presence or absence of a repeated vowel and whether or not the error outcome is a word. Table 8 gives an example of a set. Whether the error outcome is a word or not (the lexical-bias variable) can only be seen by switching the initial consonants of the pairs. In the case of *dean bead* and *dean bad,* any initial consonant ordering error, whether an exchange, anticipation, or perseveration will create one or more unintended words. For *deal beak* and *deal back* such an error will only result in one or two meaningless, but perfectly pronounceable, syllables.

Table 8
*Example of a Critical Pair Set From Experiment*

| Target | Outcome | Condition |
|---|---|---|
| Dean bead | Bean deed | Word-outcome repeated vowel |
| Dean bad | Bean dad | Word-outcome different vowel |
| Deal beak | Beal deak | Nonsense-outcome repeated vowel |
| Deal back | Beal dack | Nonsense outcome different vowel |

The pairs of each set were constructed so that they were minimally different from one another. The repeated- versus not-repeated-vowel factor was manipulated through a change in the second vowel of the pair for pairs with different vowels (/i/ to /æ/, in the example). The word versus nonword outcome resulted from a change in the final consonant, or the vowel and the final consonant. Thus, there was some degree of control over the sounds in the critical stimuli. In addition, the log frequency of the initial members of the pairs was controlled. For stimuli both designed to make words (*dean*) and nonwords (*deal*), the mean log frequencies were 1.40.

Four lists of pairs were assembled, each list having one pair from each set of critical pairs, so that each list contained 5 pairs in each condition, and each pair appeared exactly once across the four lists. In addition to the 20 critical pairs, each list contained 87 interference pairs, 63 filler pairs, and 22 noncritical cued pairs. In the make-up of each list, from 3 to 5 interference pairs preceded each critical pair and were designed so as to set up the subject to expect the initial consonants of that critical pair to occur in the reverse order. The interference pairs for a given critical pair did not contain any of the final consonants from the critical pair or any of the other pairs in its set. The filler and noncritical cued pairs were interspersed throughout the lists to prevent subjects from discovering the pattern of interference and anticipating the signal to speak, which followed each cued pair. Each list contained the same interference, filler, and noncritical cued pairs in exactly the same order. The only difference among the lists was which member of the sets of critical pairs was inserted.

*Procedure.* Word pairs were presented one at a time on a CRT screen under the control of a microcomputer system. Each pair was presented in lower case letters in the upper left part of the screen, using an 80-character/line format. Each pair was presented for 0.9 s, followed by 0.1 s of blank screen. Following the presentation of each cued pair (both critical and noncritical cued pairs), a series of question marks was presented. This was the signal for the subject to speak the pair. Then, after either 500 ms, 700 ms, or 1,000 ms, depending on which deadline condition the subject was in, a tone sounded, marking the end of the time in which the subject was given to say the pair. One-half second after the tone, the phrase DID YOU MAKE AN ERROR? was presented for 2.5 s. The subject then said "yes" aloud if he or she said anything deviant from his or her intention, and "no" otherwise. Following that, the phrase REPEAT WORDS was presented for 2.5 s, and the subject was to repeat slowly the correct word pair. Following the presentation of the command for

subjects to repeat the pair again, the screen presented an instruction for the subject to press a key on the terminal keyboard when they were ready to continue. Thus, there was a subject-controlled break after each cued pair.

Each subject received a single list under one of the three deadline conditions. Subjects were instructed to prepare to say each word pair as they saw it and to say the most recent pair aloud, clearly, upon seeing the question marks. In addition, they were told to finish before the tone sounded. During practice trials (24 pairs, 8 of which were cued), they were urged to speak faster whenever they failed to finish saying the pair within the deadline. The experimental sessions were tape-recorded.

*Subjects.* Subjects were 132 undergraduates. They were randomly assigned to lists and deadline conditions so that each list–deadline combination had 11 subjects. The subjects were volunteers from an introductory psychology course, who received extra credit for participating.

*Preliminary data analyses.* Overall, 269 initial consonant slips involving critical pairs occurred, for a 10.2% error rate. Of these, 256 were clearly errors of output, because the subjects either stated that they had erred in the judgment task or had overtly corrected themselves while they were speaking. The 13 remaining errors could have been reading errors, because for these, subjects behaved as if they had not erred—in the judgment task they said "no," and in the recall task they recalled what they said, rather than the correct stimulus pair. There were tendencies for the reading errors to occur more often in the word-outcome conditions (9 cases) than in the nonword-outcome conditions (4 cases), and to occur more often in the long-deadline condition (6 cases). Only the 256 detected errors were included in subsequent analyses.

Errors were coded as either exchanges, incomplete exchanges, anticipations, or perseverations. A complete exchange was counted when a subject reversed the order of the initial consonants (e.g., *deal back* → *beal dack*). To count as an anticipation, the initial consonant of the second word had to replace that of the first word, and the second word must be correct (e.g., *beal back*). Incomplete exchanges occurred when subjects anticipated the initial consonant of the second word, but then either stopped completely or corrected themselves before coming to the second word (e.g., *beal . . . er deal back*). Notice that errors in this category could have turned out as either exchanges or anticipations. Finally, a perseveration was counted when the first word was correct and the second word's initial consonant was replaced by the initial consonant of the first word (e.g., *deal dack*).

Some initial-consonant errors could not be clearly classified. The difficulty in coding arose when subjects stopped in the middle of syllables. The rule adopted for coding purposes was that a syllable that is begun is assumed to be completed. Thus the error *bea . . . uh deal back* is an incomplete exchange, but *beal d . . . uh deal back* is an exchange. Sometimes, however, it was hard to tell whether the second syllable of the error was begun. A second scorer independently coded 50 of the errors, and there was some disagreement involving the categories of incomplete and complete exchanges. Of the 40 incomplete and complete exchanges, the second scorer disagreed with the main scorer on 7 of them. Thus, there is some subjectivity in the coding of these two categories. Other than that there was little problem in coding. Perseverations and anticipations were easy to score because subjects did not stop in the middle of speaking. The decision as to whether an error of any type occurred was "coded' by the subjects themselves, inasmuch as the error had to be detected by them as a slip, for inclusion in the analysis.

Table 9
*Initial Consonant Error Frequencies From Experiment*

| Error type | Condition | | | | Total |
|---|---|---|---|---|---|
| | WR | WN | NR | NN | |
| **500-ms deadline** | | | | | |
| Exchange | 13 | 8 | 10 | 8 | 39 |
| Incomplete exchange | 10 | 15 | 13 | 24 | 62 |
| Anticipation | 2 | 2 | 2 | 3 | 9 |
| Perseveration | 1 | 0 | 0 | 1 | 2 |
| Total errors | 26 | 25 | 25 | 36 | 112 |
| Lates | (15) | (7) | (7) | (12) | (41) |
| **700-ms deadline** | | | | | |
| Exchange | 8 | 6 | 8 | 4 | 26 |
| Incomplete exchange | 13 | 15 | 12 | 10 | 50 |
| Anticipation | 4 | 5 | 0 | 0 | 9 |
| Perseveration | 2 | 1 | 1 | 0 | 4 |
| Total errors | 27 | 27 | 21 | 14 | 89 |
| Lates | (14) | (4) | (7) | (6) | (31) |
| **1,000-ms deadline** | | | | | |
| Exchange | 4 | 1 | 2 | 0 | 7 |
| Incomplete exchange | 11 | 14 | 9 | 10 | 44 |
| Anticipation | 2 | 0 | 0 | 0 | 2 |
| Perseveration | 1 | 1 | 0 | 0 | 2 |
| Total errors | 18 | 16 | 11 | 10 | 55 |
| Lates | (6) | (0) | (2) | (4) | (12) |

*Note.* There were 220 error opportunities in each condition. WR = word-outcome, repeated vowels; WN = word-outcome, nonrepeated vowels; NR = nonsense-outcome, repeated vowels; NN = nonsense-outcome, nonrepeated vowels.

The number of errors of each type as a function of conditions is shown in Table 9, along with the number of *late* articulations. A late articulation was an event in which the subject either failed to say anything following the cue to speak, or began to speak only after the deadline. Two kinds of analyses were performed. First, analyses of variance were done using error totals collapsed over types as dependent variables. One of these used the number of errors per condition for each subject as the dependent variable ($F_1$ statistic); the other used the number of errors per condition for each critical pair set as the dependent variable ($F_2$ statistic). The second type of analyses were nonparametric in nature and focused on the effects of specific error types. The next three subsections highlight the important findings of these analyses.

*Results involving lexical bias.* The main result of the experiment was an interaction between the lexical-bias factor (whether the error created a word or a nonsense syllable) and the deadline factor, $F_1(2, 120) = 3.94$, $p < .05$; $F_2(2, 32) = 3.30$, $p < .05$. At the 500-ms deadline, the fast speech rate, errors were

common but there was no tendency for them to make words (51 in the word-outcome condition vs. 61 in the nonword-outcome condition). Only at the longer deadlines was there a lexical-bias effect, 54–35 at the 700-ms deadline and 34–21 at the 1,000-ms deadline. So, the model's prediction that lexical bias should increase as speech slows was confirmed.

The findings of no lexical bias at the short deadline reveals something of the nature of output biases. These biases, or at least the lexical bias, are not always in force during speaking. That is, it is possible to speak and not be compelled toward saying words or wordlike syllables. According to the model, the governing factor is the time available for speaking. However, an alternate explanation for the results cannot be ruled out. This is that subjects in the fast-deadline conditions decide that there is too little time to engage in active, pre-articulatory editing of their output. They "turn off" their editorial processes and thus lose their biases toward meaningful and frequently occurring sound combinations (see Baars et al., 1975; Motley et al., 1982).

*Results involving the repeated-phoneme effect.* From the model's point of view, the results for the repeated- versus different-vowel conditions were disappointing. Pooling over error types, there was no interaction between the repeated-phoneme variable and the deadline, $F_1 < 1$, $F_2(2, 32) = 1.32$. The predicted pattern of an increasing repeated-phoneme effect as speech slows can be seen in the error totals only minimally. At the 500-ms deadline there were 51 errors in the repeated condition and 61 in the different-vowel condition, and at the two longer deadlines these numbers were 77 and 67, respectively.

One feature of the repeated-phoneme effect in previous research has been that the effect has been limited largely to exchanges, as opposed to anticipations and perseverations (Dell, 1984; MacKay, 1970). This tendency is present in the data in Table 9. For exchanges there was a repeated-phoneme effect of 45–27 ($p < .05$, by sign tests across subjects and critical pair sets), but not in the other error types (83–101 in the wrong direction, *ns*). The effect with exchanges appears to increase as predicted when the deadline was longer—23–16 at 500 ms, 16–10 at 700 ms, and 6–1 at 1,000 ms, but the numbers are too small to draw conclusions regarding an interaction with deadline.

*Results involving error types at different deadlines.* As should be apparent by now, the most potent experimental factor was the deadline, $F_1(2, 120) = 7.32$, $p < .01$; $F_2(2, 32) = 12.19$ $p < .01$, with error rate and deadline showing the expected speed–accuracy trade-off. As predicted, the trade-off was strongest for exchange errors. At the fast rate, complete exchanges made up 35% of the errors. When speech slowed, this percentage dropped. For the 700-ms deadline there were 29% complete exchanges, and for the 1,000-ms deadline only 13% of the errors were complete exchanges. There was a significant interaction between exchanges versus other errors, and the deadline condition, $\chi^2(2) = 6.85$, $p < .05$. (Note: This analysis used the subject as the unit of categorization, not the individual error.)

According to the model, an error that would be an exchange at a fast rate might only be an anticipation at a slow rate. This is why the number of exchanges was predicted to drop off most steeply with speech rate. However, the relatively fewer exchanges at the slower rates could reflect something else. When speech is slow, subjects can easily stop in the middle of what they are saying if they wish. Perhaps, there were fewer exchanges at the slow rates simply because there were relatively more incomplete exchanges, errors in which subjects stopped speaking before they got to the second syllable. If incomplete exchanges are left out, the proportion of exchanges no longer significantly differs across deadlines, although it does decrease as predicted, 78%, 67%, and 64% for the 500-ms, 700-ms, and 1,000-ms deadline, respectively.

A final point concerning the experiment's results is the main effect of error types. As seen in Table 9, complete exchanges outnumber anticipations and perseverations to a large extent. This may seem inconsistent with the fact that in natural errors, exchanges are the least likely of the three types. The experiment, however, is a special situation that is designed to produce exchanges through phonemic set expectancies. The large number of exchanges, thus, should not be seen as inconsistent with the natural data. Nevertheless, this raises the issues of exactly what is going on in experiments of this type and how they can be modeled. The next part of this section will develop the model so that it can reflect the Slip procedure and, thus, test the model by attempting to fit the results of the experiment.

### *Model of the Slip Procedure: The Slip Simulation*

The significant findings of the experiment included large main effects of deadline and error type, an interaction between lexical bias and deadline, and an effect of repeated phonemes on complete exchanges. To show that the model is consistent with these findings, I will attempt to simulate the obtained error percentages in the experimental conditions. The challenge is to *simultaneously* reproduce the various effects and their sizes with a single set of parameters, thus showing some degree of internal consistency in the way that different effects are explained. Unfortunately, the previously presented simulation cannot be used for this purpose because its vocabulary was just not set up to reflect the conditions of the experiment.

Because of this difficulty, a new, simplified simulation was constructed, which will be called the Slip simulation. The original simulation's network had many different kinds of nodes (morphemes, syllables, constituents, clusters, null elements, features). Its purpose was to show how the frequencies of the types of slips could be explained through a combination of spreading activation and phonological structure assumptions. In order to demonstrate that the model explains lexical-bias and repeated-phoneme effects, one can get by with a smaller, simpler simulation.

The Slip simulation's network contained only morphemes (16) and phonemes (4 initial consonants, 4 vowels, and 4 final consonants, each equally likely). Technically speaking, there was a different network for each of the four stimulus conditions of the experiment: the word-outcome, repeated-vowels condition (WR); word-outcome, different-vowels condition (WN), the nonword-outcome, repeated-vowels condition (NR); and the nonword-outcome, different-vowels condition (NN). Each of these four networks was the same except for a critical part whose nodes and connections were arranged to reflect the condition (see Figure 3).

For the two word-outcome conditions (WR and WN), the critical parts contained four morpheme nodes and their connections to phonemes. Two of these morphemes represented the two target words of the critical pair. These nodes are labeled as

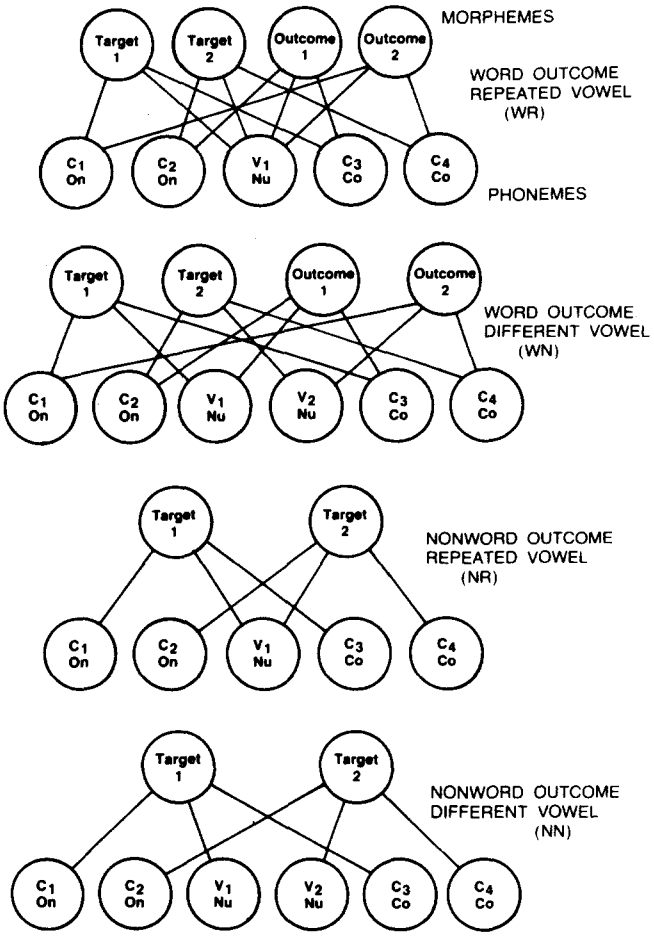


*Figure 3.* The critical parts of the network used in the Slip simulation. (For each condition, a different critical part was embedded in the network. Each (C)onsonant and (V)owel node is labeled as to whether it is a potential (On)set, (Nu)cleus, or (Co)da. All connections are top-down and bottom-up.)

*Target 1* and *Target 2* in Figure 3 and will be referred to by these labels. The other two morpheme nodes in the critical section of the networks of the word-outcome conditions represented the potential word outcomes created by exchanging the initial consonants of Target 1 and Target 2. (These are labeled *Outcome 1* and *Outcome 2*.) For the WR condition, Target 1, Target 2, Outcome 1, and Outcome 2 connected to the same vowel node. For the WN condition, Target 1 and Outcome 1 connected to one vowel and Target 2 and Outcome 2 connected to another vowel.

The critical parts of the networks for the nonword-outcome conditions (NR and NN) were identical to their respective word-outcome conditions (WR and WN) except that there were no outcome nodes. In these nonword-outcome conditions the exchanges of the initial consonants of the intended morphemes does not result in morphemes. So the networks for these conditions had no nodes representing the outcomes.

The basic idea of the Slip simulation is to simulate the utterance of critical pairs in the four experimental conditions under the three experimentally manipulated speech rates. So, in order to simulate the utterance of a critical pair in the WR condition under the 500-ms deadline, the network with the WR critical part would be used. The simulation would attempt to encode Target 1 followed by Target 2, and a certain number of time steps (500 ms worth) would be used. Other experimental conditions would be simulated using the network with the appropriate critical part embedded in it to represent different critical pair types, and a different number of time steps to represent different deadlines.

The Slip simulation worked largely the same as the previous one. To encode a critical pair, the first intended morpheme, Target 1, was given an arbitrary amount of activation (60 units) and the second intended morpheme, Target 2, was primed with one half of that amount (30 units). Activation then spread according to the rule given earlier. Each node sent some fraction of its activation to all of the nodes directly connected to it during a time step, and each node lost a certain fraction of its activation during a time step. The spreading continued for $r$ time steps, where $r$ varied with the deadline condition, and was determined by a parameter, $t$, that specified how long in real time each time step was. After the $r$ steps passed, the most highly activated initial consonant, vowel, and final consonant were selected, thus encoding the first morpheme of the pair. The activation levels of the selected phonemes were set to zero, and the encoding of the second morpheme began. Its node was given 60 more units of activation, and the process continued for $r$ more time steps, culminating in the selection of the initial consonant, vowel, and final consonant of the second morpheme.

At this point it becomes important to say how the simulation was constrained in fitting the data. Certain features of the simulation were set up in advance, such as the structure of its networks and the main features of its operations. These were derived from the phonological encoding model and the stimulus conditions of the experiment. Other aspects of the simulation, its free parameters, were permitted to vary during the fitting process. One feature that was set up in advance, however, was somewhat like a free parameter. This was the fact that when the first morpheme is activated with 60 units, the second one is primed with exactly one half of that, 30 units. Conceivably, it could be primed by some other fraction, and this might affect the model's performance. Thus, this ratio of 30/60 was a hidden parameter. It was not a free parameter, technically, because it was not varied, but it had no claim to any motivation other than my belief that it would be a good value.

The simulation had five free parameters, the spreading rate $p$, the decay rate $q$, the time parameter $t$, and two parameters designed to reflect the interference pairs in the Slip procedure, $\mu_1$ and $\mu_2$.

With regard to the spreading rate, it was assumed that each connection, both those from morpheme to phoneme and those from phoneme to morpheme, had the same spreading rate. This rate, $p$, was the fraction of activation that was sent in each time step. In addition, the fraction of its activation that each node lost during a time step, $q$, was assumed to be the same for all nodes.

The time parameter $t$ specified how long a time step was. In so doing, it determined how many time steps, $r$, should pass during the encoding of each syllable under each deadline condition. To keep things simple in translating from $t$ to $r$, I assumed for the simulation that the first morpheme was always spoken halfway through the deadline and the second morpheme was

spoken at the end of the deadline. Thus, for the 500-ms deadline, 250 ms worth of time steps passed before the phonemes of the first morpheme were selected, and 500 ms worth passed before those of the second morpheme were selected. I assumed that the sounds of a morpheme were selected simultaneously; that is, no time steps passed between the selection of the three phonemes.

In order to characterize the final two free parameters, $\mu_1$ and $\mu_2$, I have to describe how the effect of the interference pairs was modeled in the Slip simulation, and give some details of how the data were fitted. In order for the simulation to represent the procedure used in the experiment, it had to include the order confusion that is created by the use of the interference pairs that prededed each critical pair. The interference pairs worked by creating a "set" for subjects to say the initial consonant of the second word first and the initial consonant of the first word second. So if the critical pair was *dean bead* the interference would set up the subject to expect a word beginning with /b/ followed by a word beginning with /d/. To model this expectation, I added a random amount of activation, $X_1$, to the initial phoneme of Target 2 (to /b/ in the preceding example) when Target 1 was beginning to be encoded. Then, when it came time to encode Target 2, a random amount of activation $X_2$ was added to the initial consonant of Target 1, that is, to /d/ in the example. Thus, $X_1$ represents anticipatory bias and $X_2$, perseveratory bias. Although this is an unsophisticated approach to the problem of expectancy, it does capture, in activation terms, what the interference does. It creates uncertainty regarding the order of the initial consonants, which, in turn, leads to slips. Parameters $\mu_1$ and $\mu_2$ characterize the distribution of the random variables $X_1$ and $X_2$. Before I discuss the nature of these distributions of $X_1$ and $X_2$, I have to say something about how the other parameters were varied to fit the data.

The fitting process consisted of two stages. In the first stage the simulation encoded Target 1 followed by Target 2 for the four stimulus conditions, with varying values of $p$, $q$, $r$, $X_1$, and $X_2$. Whether the result was a correct encoding, an exchange, an anticipation, or a perseveration of the initial consonants was recorded. The second stage used the results of the first and determined the probability distributions of $X_1$ and $X_2$ so that the simulated error probabilities were as similar as possible to those of the experiment. I should add that the whole process was largely done by trial and error and that the final "best fit" was just the best fit that I could find. There was no formal search of the parameter space to find a fit that minimized the sum of squares or chi-square.

During the first stage of fitting it was useful to generate what I call *error profiles*. An error profile is the set of outcomes that resulted from the simulation encoding the critical pair in a given stimulus condition (WR, WN, NR, or NN), with constant values of $p$, $q$, and $r$, and with varying values of $X_1$ and $X_2$, the interference random variables. The values of $X_1$, and $X_2$ that were looked at ranged from 0 to 40, by integral steps. Graphically, a profile can be seen as a matrix in which columns represent values of $X_1$ and in which rows represent values of $X_2$. Figure 4 presents such a matrix for the profile for the WR condition, with $p = .1$, $q = .3$, and $r = 3$ time steps per syllable. Each entry in the matrix presents what would happen with that particular combination of $X_1$ and $X_2$—a C means that no error occurs, an A means an anticipation occurs, P means a perseveration, and $E$ means an exchange. Consider the entry for $X_1 = 2$, and $X_2 = 3$. This means that two units of activation were added to the initial consonant of Target 2 when Target 1 was beginning to be encoded and three units were added to the initial consonant of Target 1 when Target 2 was due to be encoded. The entry in this cell, C, means that there was no error under these conditions. The values of $X_1$ and $X_2$ were not extreme enough to cause the simulation to err.

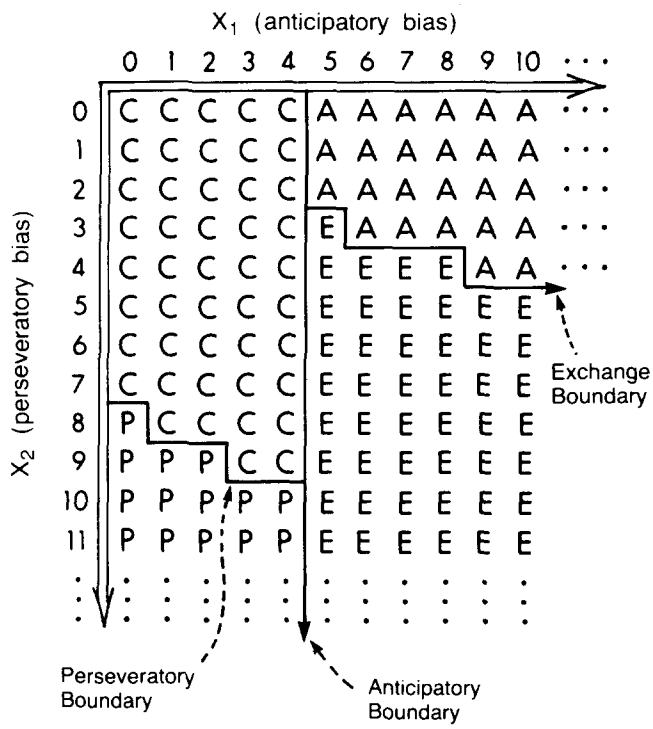


*Figure 4.* A sample profile of outcomes from the Slip simulation for the word-outcome, repeated-vowel condition with $p = .1$, $q = .3$, and $r = 3$ time steps per syllable. (C, A, E, and P represent, respectively, correct outcome, anticipation, exchange, and perseveration.)

The profiles are interesting in themselves, in addition to their usefulness in fitting. Each contains three boundaries, an anticipatory boundary, an exchange boundary, and a perseveration boundary. The anticipatory boundary (the vertical line between $X_1 = 4$, and $X_1 = 5$), is the minimum value of $X_1$, the anticipatory bias, required to get an anticipatory slip, either an anticipation or an exchange. That is, we need at least five units of activation added to /b/ in order to make the slip *dean bead* → *bean*. . . . Whether this slip turns out to be an exchange *bean deed*, or an anticipation *bean bead*, depends on both $X_1$ and $X_2$. What I call the exchange boundary is the border between the anticipation and exchange portions of the profile. The fact that this border is sloped downward simply reveals the fact that when $X_1$, anticipatory bias, is large, relatively more perseveratory bias, $X_2$, is needed to make the slip into an exchange. Notice that the exchange boundary is "higher" than the boundary between the correct encodings and the perseverations (the perseveration boundary). This is a graphic way of saying that in the model, an exchange is not simply the independent occurrence of an anticipation and perseveration involving the same words.

Table 10
*Fit of Slip Simulation to Obtained Error Percentages From Experiment*

| Error type and condition | Deadline | | |
|---|---|---|---|
| | 500 ms | 700 ms | 1,000 ms |
| Exchange | | | |
| WR | 9.6 (11.9) | 8.0 (7.4) | 5.7 (7.0) |
| WN | 9.2 (11.9) | 7.7 (7.2) | 5.4 (6.3) |
| NR | 9.3 (11.9) | 7.7 (6.6) | 4.1 (2.4) |
| NN | 12.5 (11.9) | 5.2 (4.1) | 3.5 (1.3) |
| Anticipation | | | |
| WR | 1.9 (1.6) | 3.3 (1.9) | 2.0 (1.3) |
| WN | 2.2 (1.6) | 4.0 (2.2) | 1.4 (2.0) |
| NR | 2.0 (1.6) | 1.4 (2.7) | 0.9 (2.4) |
| NN | 3.4 (1.6) | 1.2 (2.4) | 1.0 (1.3) |
| Perseveration | | | |
| WR | 0.5 (0.05) | 0.9 (<0.01) | 0.5 (<0.01) |
| WN | 0.0 (0.05) | 0.5 (<0.01) | 0.5 (<0.01) |
| NR | 0.0 (0.05) | 0.5 (<0.01) | 0.0 (<0.01) |
| NN | 0.5 (0.05) | 0.0 (<0.01) | 0.0 (<0.01) |

*Note.* Obtained percentages are followed by simulated percentages in parentheses. The simulation used $p = .3$, $q = .4$, t (one time step) = 250 ms, $\mu_1 = 18$, $\mu_2 = 10$. The predicted data for the 700-ms deadline were derived from a linear interpolation of predictions using 2 time steps (500 ms) and 3 times steps (750) ms. The obtained data are derived from that in Table 9 as described in the text. WR = word-outcome, repeated vowels; WN = word-outcome, nonrepeated vowels; NR = nonsense-outcome, repeated vowels; NN = nonsense-outcome, nonrepeated vowels.



Rather, exchanges can happen with less perseveratory bias than is necessary to make a perseveration.

Changing the stimulus conditions (e.g., from WR to WN), or changing the speech rate (e.g., from $r = 3$, to $r = 6$), changes the boundaries of each profile in predictable ways. Slowing down the speech rate by increasing $r$ moves the anticipatory boundary to the right and the exchange and perseveration boundaries downward. The net result is that the "correct" area in the upper left corner of the profiles is larger. These shifts in the boundaries show graphically that the model's error rate should decrease as speech slows, because it takes more bias, that is, more extreme values of $X_1$ and $X_2$, to cause slips at the slower rates. More interesting, the boundary shifts associated with rate changes reveal the model's prediction of an interaction between rate and error types. Moving the anticipatory boundary to the right and lowering the other two boundaries reduces the region in the profile for exchanges. However, these changes do not necessarily reduce the regions for anticipations and perseverations. The anticipation region is reduced because of the rightward shift of the anticipatory boundary, but it is increased by the downward shift of the exchange boundary. Similarly, the perseveration area is reduced by the downward shift in the perseveration boundary but increased by the leftward shift of the anticipatory boundary. Whether the model makes more or fewer anticipations and perseverations when speech slows down depends on the probabilities of each $X_1X_2$ combination. But regardless of these probabilities we can say that when speech slows, the chance of an exchange is decreased by a greater factor than that of an anticipation or a perseveration.

Similar boundary shifts occur when the lexical-bias and repeated-phoneme variables are considered. Comparing profiles for word-outcome with nonword-outcome conditions, with all other conditions and parameters held constant, shows that the correct region is larger for the nonword conditions. This illustrates the model's prediction of a lexical-bias effect. By the same token, comparing profiles for repeated- and not-repeated-phoneme conditions shows the correct area to be larger for the not-repeated-phoneme conditions. Generally speaking, the less error-prone conditions, the nonword-outcomes and the different-phoneme conditions, show right-shifted anticipatory boundaries and downward-shifted exchange and perseveration boundaries. For example, consider the anticipatory and exchange boundaries for the conditions when $p = .3$, $q = .4$, and $r = 4$. For the word-outcome conditions the average location of the anticipatory boundary is at $X_1 = 15$, and for nonword outcomes, at $X_1 = 17.5$. Thus, a potential word outcome requires less anticipatory bias to slip. For repeated-phoneme conditions, the anticipatory boundary is at $X_1 = 16$, and for different-phoneme conditions it is at $X_1 = 16.5$. In this case, and generally as well, the model's lexical-bias effect is stronger than its repeated-phoneme effect. Next consider the exchange boundaries. For nonword conditions it is four units lower than for word-outcome conditions, and for different-phoneme conditions it is one unit lower than for repeated-phoneme conditions.

These aspects of the profiles reveal two unsuspected predictions from the model. Recall that the rightward shift of the anticipatory boundary and the downward shift of the other two boundaries, all of which occurred as the speech rate slowed, implied a relatively greater reduction in the exchanges than in the other two error types. Because the shift pattern is similar for the lexical-bias and repeated-phoneme variables, these variables should also interact with error types. In particular, the repeated-phoneme and the lexical-bias effects should, like the speech-rate effect, be more evident in exchanges than in the other types.

Is there any evidence for these predicted interactions? I have shown elsewhere (Dell, 1984) that the repeated-phoneme effect is more evident for exchanges than for anticipations and perservations in both errors generated in the Slip procedure and in sound errors from error collections. The experiment presented here only showed a weak repeated-phoneme effect, but in support of the prediction, it was significant only for the complete

exchanges. Thus, the predicted interaction of error type with repeated phoneme has some support. The other prediction, that the lexical-bias effect should be more evident with exchanges, has more checkered support. Baars et al. (1975) found that the tendency for slips to make words was stronger with complete exchanges than with the other error types. However, this interaction was not replicated in the present experiment. One can see in Table 9 that the greater number of word-outcome over nonword-outcome slips does not differ among the error types. Where there is lexical bias, at the 700-ms and 1,000-ms deadlines, it is present for all error types. Thus this experiment, in contrast with Baars et al., did not verify the model's prediction that lexical bias should be stronger in exchanges than in anticipations or perseverations. If present, the effect is not a robust one.

I have already explained why the model predicts that speech-rate effects should be more pronounced with exchanges. Why should the lexical-bias and repeated-phoneme effects also be stronger with exchanges? The reason is that an exchange requires two erroneous phoneme selections, one in each word. The other error types are characterized by only a single misselection. Thus, any factor that contributes to a misselection, whether a fast speech rate, a repeated sound, or the fact that the outcome string is a word, is going to be doubly important in causing an exchange. This can be seen in the profiles. When the exchange area expands due to the action of some variable, two factors are at work. The anticipatory boundary moves to the left, and the exchange boundary moves up. These two boundary shifts can be seen as reflecting the double nature of exchanges, the fact that they involve two substitutions.

As we have seen, the profiles do not, themselves, specify the probabilities of the simulation's error types. To do this, one must consider the probability of each $X_1X_2$ combination. This consideration governed the second stage of the fitting process.

In order to characterize the probability distribution of the $X_1X_2$s, I assumed that low and moderate values of bias would be more probable than high ones and that bias would never be negative. Both of these assumptions are reasonable in light of what $X_1$ and $X_2$ represent, namely, the degree of bias created by the interference pairs in the Slip procedure. Given these assumptions and some practical considerations, I chose to characterize $X_1$ and $X_2$ by independent Poisson distributions. The Poisson distribution is a discrete probability function whose minimum is zero. It can be specified by a single parameter, $\mu$, which is both its mean and variance. Thus, it was assumed that $X_1$ was sampled from a Poisson distribution with mean $\mu_1$, and $X_2$ was sampled from a like distribution with mean $\mu_2$. This characterization of the probabilities of the anticipatory and perseveratory bias is really one of convenience more than one of principle. It was convenient to use single parameter distributions and to make them independent. In the same vein, it was convenient to use discrete functions because the profiles were, themselves, generated using discrete values of $X_1$ and $X_2$.

In order to fit the experimental data, I tried out a number of values of $\mu_1$ and $\mu_2$ on a large number of profiles and found a parameter combination that gave good results. The parameters, and the predicted and obtained error probabilities, are given in Table 10. Before discussing and evaluating the fit, I should point out that the obtained error probabilities have been changed so that there is no longer an incomplete exchange category. I assumed that the errors designated as incomplete exchanges were, in fact, exchanges and anticipations. The incomplete exchanges were placed in the exchange and anticipation categories in such a way that the ratio of exchanges to anticipations at each deadline was not changed by the influx of errors from the incomplete exchange category. Thus, about 80% of the incomplete exchanges (the percentage varies with the deadline) were assumed to be exchanges and the remainder, anticipations.

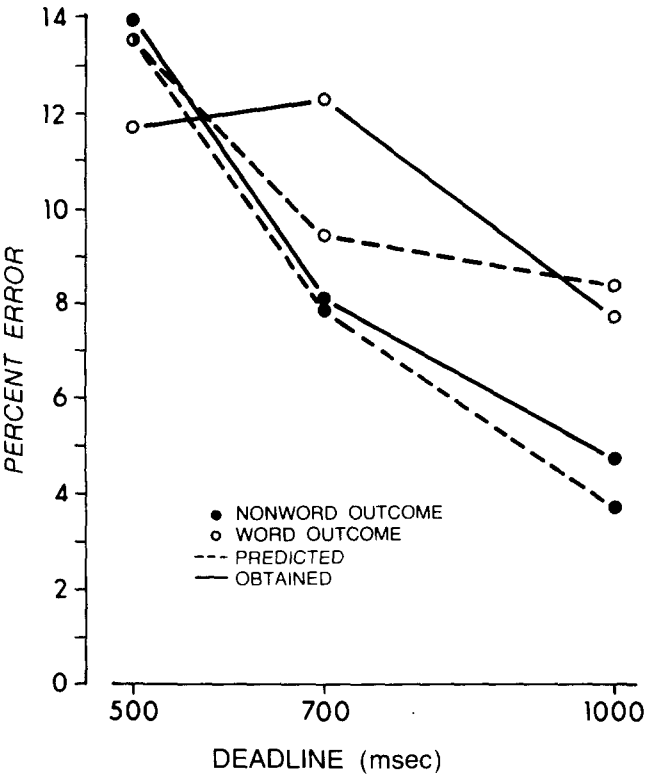


*Figure 5.* Initial-consonant error rate (all errors) as a function of word versus nonword outcome and the deadline. (Obtained data are from the experiment, and predicted data are from the fit of the Slip simulation described in Table 10.)

How well does the model reflect the data? Figures 5 through 9 highlight aspects of the fit by presenting obtained and predicted curves for some of the key interactions among the variables—lexical bias with rate (Figure 5), repeated phoneme with rate (Figure 6), error types with rate (Figure 7), lexical bias with error types (Figure 8), and repeated phonemes with error types (Figure 9). With the exception of the predicted interaction between lexical bias and error types, the model is accounting for the major patterns in the data. Given that these interactions and even the main effects are emergent properties of the model, not tied to specific parameters, the fit can be thought of as adequate. One worrisome point concerns the perseverations. These errors occurred very infrequently, and with the chosen set of parameters they should occur infrequently. But at the longer deadlines, their predicted frequency was, for all practical purposes, zero. Yet, they did occur. Thus, the model's account of perseverations in the Slip procedure could be flawed.

A final point regarding the fitting concerns the particular networks used. Recall that the critical part of each network (the

part that contained the defining nodes and connections for each stimulus condition) was embedded in a larger network that was essentially a random collection of additional morphemes, phonemes, and their connections. To show that these extra nodes have little impact on the behavior of the model, I generated two additional sets of extra nodes and connections in which to embed the critical sections. Using these new networks and the same parameter values, the predicted error proportions were nearly the same as with the original one. Table 11 presents the predicted proportions for the original (Net 1) and the two new networks (Nets 2 and 3). One can conclude that the model's performance with this set of parameters is determined much more by the local region containing the intended morphemes (the critical part) than by the extra nodes outside this region.

To conclude the discussion of the Slip simulation, I want to stress that it is not exactly the same theoretical entity as the phonological encoding model that provided the basis for the first, larger simulation. Rather, it was derived from the phonological encoding model and included additional assumptions that made it suited to the experimental procedure. A similar relation holds between the phonological encoding model and the general theory of production. The phonological encoding model was derived from the theory, but was tailored for phonological encoding. With this in mind I will finish the discussion of phonological encoding by stepping outside of both simulation models and returning to the level of the general theory.

### *Semantic and Syntactic Effects in Phonological Encoding*

So far, phonological encoding has been seen as an isolated process that maps from a morphological to a phonological representation. The errors that occur in this process, the sound errors, were seen to be influenced by phonological variables such as the structure of syllables and speech sounds, and by the items present in the mental lexicon. Any complete theory of phonologial encoding, however, must recognize that it takes place against a background provided by other higher level processes that influence it in various ways. Some of these influences have been crystallized as syntactic and pragmatic conditions on phonological rules (e.g., Cooper & Paccia-Cooper, 1980; Gazdar, 1980). So, for example, *want to* can be contrasted to *wanna* in certain syntactic contexts when speech is informal. Other higher level influences emerge as syntactic and semantic influences on the probability of sound errors. The general theory of production provides a natural mechanism for these latter influences because higher level encoding processes, which are going on simultaneously with phonological encoding, are indirectly, but continually affecting the activation levels of phonological nodes. As a result, the theory makes predictions regarding semantic and syntactic effects on sound errors. Although these predictions have not been quantitatively explored through simulation, their existence can be deduced from the theory.

First, consider a semantic effect. When a particular word is being phonologically encoded, other word and morpheme nodes are activated for various reasons. Some of them are active because they share components of meaning or are associated with other words in the sentence. For example, in the sentence *The doctor has a new purse,* at the time of encoding *purse* we would expect *purse* to be highly active, and *new* and *doctor* to be somewhat less active. In addition, the word *nurse,* being associated with *doctor* and sharing sounds with *purse,* should be active. This state of affairs would, in the theory, set the stage for the perseveration of /n/ in *new* to create *The doctor has a new nurse.* The fact that *nurse* is an existing morpheme helps make the slip more likely (lexical bias), and the semantic association to *doctor* increases this likelihood. Earlier, this phenomenon was referred to as semantic bias, and it was mentioned that there is some evidence for such a bias in errors (Fromkin, 1971; Motley & Baars, 1976).

Next, let us consider some predicted syntactic effects on sound errors. In sound misorderings, a sound or group of sounds moves from one syllable to another. It has often been noted that this movement tends to occur within syntactically defined chunks—major phrases or clauses. For example, Garrett (1975) reported that only 7.5% of sound exchanges from his corpus crossed clause boundaries. The syntactic constraint on movement suggests that sounds within a syntactic chunk have a greater chance of being simultaneously active than those from different chunks. This suggestion fits nicely with work by Cooper and colleagues (e.g., Cooper, Egido, & Paccia, 1978) showing that phonological adjustments such as *did you* → /ju/ tend not to occur when the two words are in separate clauses. Other evidence from pauses and other measures of processing load during spontaneous speech point to a syntactic construction process that proceeds in a clause-by-clause fashion (e.g., Ford, 1978; Ford & Holmes, 1978; Grosjean, Grosjean, & Lane,

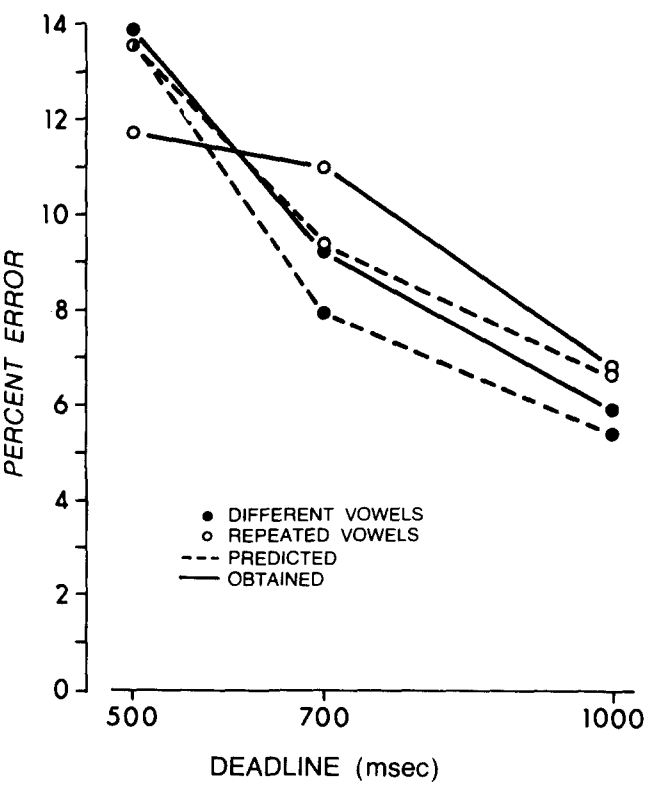


*Figure 6.* Initial-consonant error rate (all errors) as a function of repeated versus different-vowel condition and the deadline. (Obtained data are from the experiment, and predicted data are from the Slip simulation described in Table 10.)

1979). Clearly, the production theory has to assume that words are selected for the syntactic representation in fits and starts that are associated with syntactic groups of some form or another. (I will leave open the question of exactly what these groups are and whether the effects are due to the temporal characteristics of semantic processing, syntax-internal features, or both.)

Adopting such a view for the theory leads to a prediction: Anticipatory errors (anticipations, exchanges, and anticipatory additions) should exhibit a stricter syntactic constraint in movement distance than should perseveratory errors (perseverations, perseveratory additions). The tendency for sounds to be anticipated should reflect higher level representations; that is, an upcoming sound will be activated to the extent that the syntax and morphology have selected, or are selecting its mother lexical item.[2] Assuming that lexical items are selected in chunks corresponding roughly to major syntactic groups, it follows that upcoming activated sounds tend to belong to the same chunk. The case is different with regard to sounds that belong to syllables that have already been encoded. These remain activated simply because they and their neighboring nodes have not yet decayed. Thus, distance effects in perseveratory errors reflect more the rate of operation of the phonology and the decay rate, not higher level rule systems. It follows that activated sounds from already encoded syllables will sometimes belong to different syntactic chunks than the syllable currently being encoded. Thus, perseveratory movement should be much less constrained by syntactic chunks than should anticipatory movement.

Do the various sound-error types show differing syntactic distance effects? There is evidence that the interacting sounds are closer in exchanges than in anticipations (Nooteboom & Cohen, 1975) and that the interacting sounds of anticipations and perseverations as a class cross clause boundaries more often than those in exchanges (Garrett, 1980a). The first of these results is quite consistent with the theory. Distance between interacting elements is correlated with the time that occurs between their selection, and thus greater distance should favor anticipations at the expense of exchanges for the same reason that a slow rate favors anticipations. The second result is not inconsistent with the theory because in the theory, exchanges should be closer than anticipations, and anticipations should show more of a syntactic constraint than perseverations. However, the key

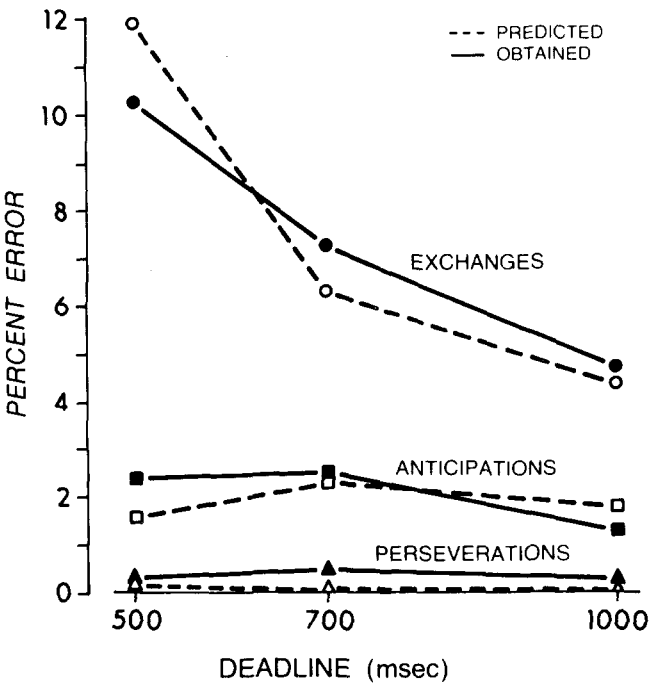


*Figure 7.* Initial-consonant error rate as a function of error type and the deadline. (Obtained data are from the experiment, and predicted data are from the Slip simulation described in Table 10.)

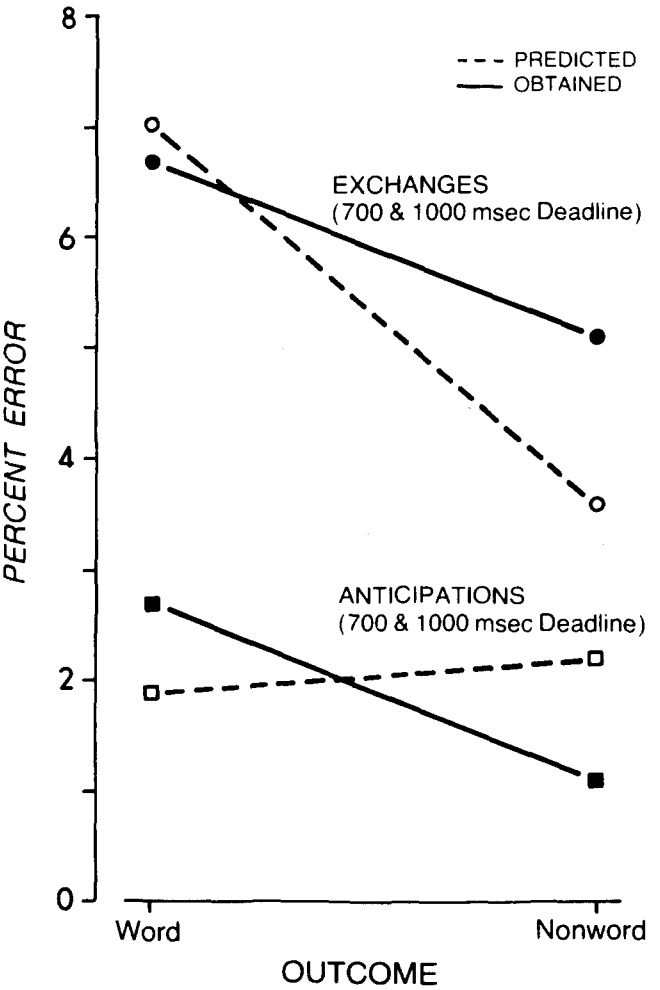


*Figure 8.* Initial-consonant error rate as a function of error type and word versus nonword outcome for errors in the 700-ms and 1,000-ms deadline conditions. (Obtained data are from the experiment, and predicted data are from the Slip simulation described in Table 10.)

[2] So far, nothing has been said about the temporal characteristics of the construction of the morphological representation, which should exert more of an influence on what sounds are simultaneously active than does the syntactic representation. There is reason to believe that the temporal characteristics of this representation may be associated with prosodic groups, which are very similar to, but distinct from, syntactic groups (Boomer & Laver, 1968; Garrett, 1982).

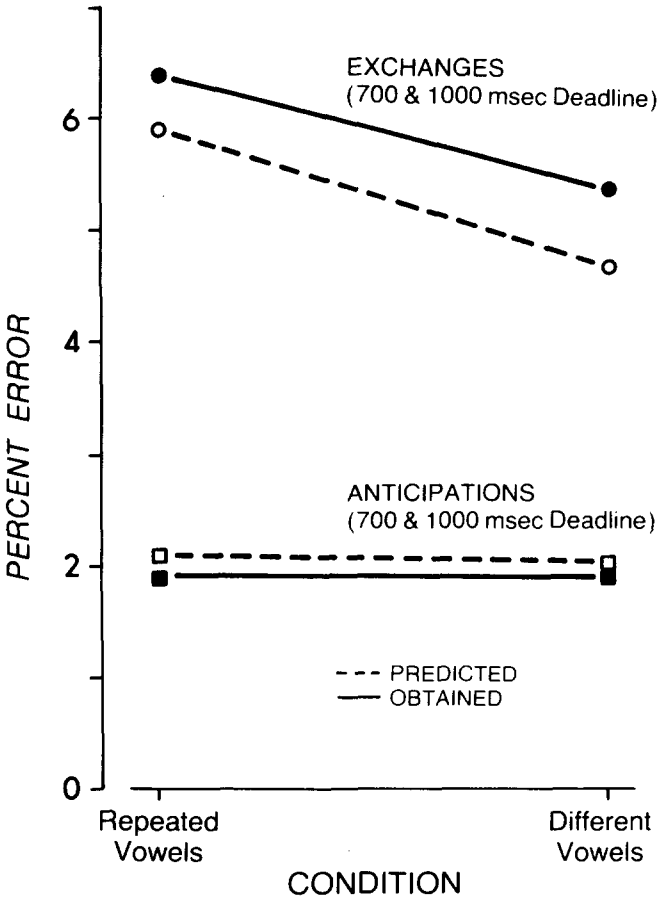


*Figure 9.* Initial consonant error rate as a function of error type and repeated- versus different-vowel condition for errors in the 700-ms and 1,000-ms deadline conditions. (Obtained data are from the experiment, and predicted data are from the Slip simulation described in Table 10.)



contrast, a comparison of the syntactic constraints on perseverations and anticipations, has not been directly examined empirically, so we remain in the dark about this prediction from the theory.

In general, the theory's account of sound-error types has the feature that the types are predicted to differ in their sensitivity to various influences (speech rate, syntactic structure, distance, lexical bias, repeated phonemes) despite their occurrence at the same level of processing involving the same mechanisms. This approach differs from that of Garrett (1980a), who argued that differences between exchanges, on the one hand, and anticipations and perseverations, on the other, suggest different mechanisms or processing levels for the types. Deciding between Garrett's proposal and the present theory is not possible with the available data, but the issue is, nonetheless, an empirical one, and one in which the present theory offers some testable predictions.

### *Some Problems in the Model*

On the whole, the model accounts for many facts about sound errors, particularly the facts that were revealed in the experiment. However, as Garrett (1980a) concluded, "error patterns as a source of data are nearly as diverse as that of normally nonerrorful structure" (p. 217), and thus it is not surprising that there are many aspects of sound errors that remain unexplained by the model. The following are five error phenomena that are not accounted for by the model as it stands and that any large scale account of sound errors really ought to deal with. (I will also suggest how the model might be changed to explain these phenomena, but it should be recognized that these changes are ad hoc.)

1. Initialness effect. The initial sounds of words and syllables tend to slip more than the other parts (e.g., MacKay, 1970). This effect is a fairly strong one. According to Stemberger (1982a) about 80% of errors involving consonants occur in syllable onsets rather than in codas. Stemberger noted that some of this difference may be due to a perceptual bias on the part of error collectors because it is easier to detect errors in initial positions (Cole, Jakimik, & Cooper, 1978), but this cannot be the whole story because the effect shows up in experimentally elicited errors (Shattuck-Hufnagel, 1981). This effect is not predicted in the model. The large simulation generated errors in onsets, nuclei, and codas about equally often (the nucleus and coda did tend to slip as a unit much more than did the onset and the nucleus, though, due to the presence of the rime constituent node).

I suspect that initial sounds slip a lot because they are, in general, easy to retrieve—or, to use activation terms, they become highly activated quickly (Brown & McNeill, 1966; Nooteboom, 1969). Thus, although the correct initial sound tends to be highly activated, so do the initial sounds of competing syllables from other parts of the utterance. As a result, these highly active competitors often replace the correct sounds. The model can be modified to reflect this explanation. If it is assumed that the connection strength (parameter $p$) from higher units to initial consonants is greater than that to noninitial sounds and that this increased strength is also associated with an increased variability in the activation levels of nodes for initial sounds, then the initialness effect on errors will emerge. However, in no way can it be said that the effect emerges directly from the model's linguistic or processing assumptions.

2. Triggering effects. Some sound errors seem to have a *trigger,* a nearby sound that is identical to the slipping target sound. For example, in the deletion error *piano sonata number* → *piano sonata umber,* the earlier /n/ in sonata, and possibly the one in *piano* as well, seem to trigger the deletion of /n/ in *number.* If this kind of triggering does occur (and some evidence indicates that it does; MacKay, 1969; Shattuck-Hufnagel, 1979) then the model is challenged. According to the model, the deletion error cited would occur when the initial null element has a higher activation level than initial /n/ when the syllable *num* is encoded. However, the preceding /n/s should tend to prevent, rather than contribute to this deletion. Repeating a sound tends to keep that sound highly activated. There is, in the model, no way that repeating a sound can weaken that sound, unless it is assumed that postselection negative feedback (the act of setting the activation level of each selected node to zero) is noisy and sometimes results in the selected node's activation being less than zero, so that the phoneme node is actually inhibited. In general, some form of inhibition is required to explain triggering effects, and thus the model must be regarded as incomplete without it.

Table 11
*Predicted Error Percentages From Slip Simulation for Three Different Networks*

| Error type and condition | Deadline: 500 ms Network 1 | 500 ms Network 2 | 500 ms Network 3 | 700 ms Network 1 | 700 ms Network 2 | 700 ms Network 3 | 1,000 ms Network 1 | 1,000 ms Network 2 | 1,000 ms Network 3 |
|---|---|---|---|---|---|---|---|---|---|
| Exchange | | | | | | | | | |
| WR | 11.9 | 11.9 | 11.9 | 7.4 | 7.4 | 7.7 | 7.0 | 7.0 | 7.6 |
| WN | 11.9 | 11.9 | 11.9 | 7.2 | 7.2 | 7.4 | 6.3 | 6.3 | 7.0 |
| NR | 11.9 | 11.9 | 11.9 | 6.6 | 6.6 | 6.6 | 2.4 | 2.2 | 2.7 |
| NN | 11.9 | 11.9 | 11.9 | 4.1 | 4.1 | 4.3 | 1.3 | 0.3 | 0.7 |
| Anticipation | | | | | | | | | |
| WR | 1.6 | 1.6 | 1.6 | 1.9 | 1.9 | 1.6 | 1.3 | 1.3 | 0.7 |
| WN | 1.6 | 1.6 | 1.6 | 2.2 | 2.2 | 1.9 | 2.0 | 2.0 | 1.3 |
| NR | 1.6 | 1.6 | 1.6 | 2.7 | 2.7 | 2.7 | 2.4 | 2.6 | 2.2 |
| NN | 1.6 | 1.6 | 1.6 | 2.4 | 2.4 | 2.3 | 1.3 | 2.2 | 2.0 |
| Perseveration | | | | | | | | | |
| WR | 0.05 | 0.05 | 0.05 | <.01 | <.01 | <.01 | <.01 | <.01 | <.01 |
| WN | 0.05 | 0.05 | 0.05 | <.01 | <.01 | <.01 | <.01 | <.01 | <.01 |
| NR | 0.05 | 0.05 | 0.05 | <.01 | <.01 | <.01 | <.01 | <.01 | <.01 |
| NN | 0.05 | 0.05 | 0.05 | <.01 | <.01 | <.01 | <.01 | <.01 | <.01 |

*Note.* All numbers are percentages. Network 1 was used to fit the obtained error data. Parameter values are given in Table 10. WR = word-outcome, nonrepeated vowels; WN = word-outcome, nonrepeated vowels; NR = nonsense-outcome, repeated vowels; NN = nonsense-outcome, nonrepeated vowels.



3. Effects of vocabulary types. Sound errors tend mostly to involve content rather than function words (Garrett, 1975; Stemberger, 1984b). It could be that this difference is due to the very high frequency of function words, their usually unstressed status, or some aspect of their syntactic function. Regardless of which of these is right, the model provides no account of the effect because it treats content and function words alike during phonological encoding. If the frequency explanation turns out to be correct, it will probably be easy to produce the effect through frequency-sensitive resting levels in activation levels (e.g., McClelland & Rumelhart, 1981). Otherwise, it is not clear how to include the effect within the context of the model.

4. Violations of the syllable position constraint. The syllable position constraint on sound errors—the rule that a misordered sound must retain its syllabic serial position—is handled by the model in a fashion that allows no violation of the constraint. The phonology generates categorized slots in the order onset, nucleus, coda, and it fills each with the most activated node of the appropriate category. Thus, an initial consonant that makes up or is part of an onset can never move to the coda position, nor can a coda move to an initial position. This treatment has two weaknesses, one empirical and one theoretical. The empirical point is that there are (rare) violations of the position constraint. Sometimes consonants move to different positions. What is more, the conditions under which violations occur seem to be somewhat predictable. According to Stemberger (1982b), MacKay (1970), and Shattuck-Hufnagel (1983), the violations almost never occur in between-word sound errors. But they do occur in within-word sound errors (e.g., *teach* → *cheach*). Stemberger found that 21.4% of the within-word consonant errors violated the constraint, whereas only 1.6% of the between-word consonant errors did. Given that these violations are predictable and not just some random noise in the system, they become noteworthy exceptions to the model's predictions.

The second major weakness of the model's account of syllable position constraint is that it necessitates different nodes for sound units in initial positions from those in final position. So, in the model, the /k/s in *cat* and *tack* and their constituent features have to be different nodes. The problem is that, from the model's perspective, there is no sense in which both sounds are /k/.

One solution to both the empirical and theoretical problems is to eliminate the position encoding of the feature nodes, but retain it for phoneme and cluster nodes, the potential onsets and codas. Then the onset /k/ and the coda /k/ would be different, but both would connect to the same set of features. This would change the properties of the model only in that it would create a slight tendency for violations of the syllable position constraint, which as we saw earlier is exactly what is needed empirically. I suspect in addition that violations would be more likely within rather than between words because the sounds within a word would tend to be co-activated to a greater extent than those from more than one word. Thus, *teach* might slip to *cheach* because the coda /č/ would be active when the beginning of the word is selected. The coda /č/ would activate the onset /č/ via their common feature nodes, thus contributing to the selection of /č/ over /t/. However, because this change in the model has not been explored by simulation, I cannot confidently assert that it has exactly these effects.

A second solution to the syllable position violations is to eliminate the position encoding of consonants entirely (returning to a single node for each phoneme) and adding special position-binding nodes that bind phonemes to syllable positions. This solution has been suggested by Stemberger (1982a) and may be workable if the binding mechanism can be specified. There would probably have to be one binding node for each possible phoneme in each position (see, e.g., Shastri & Feldman, 1984).

Regardless of how the model is modified, it is clear that it

must be modified. The model's complete separation of initial from final consonants fails to explain the sense in which the initial and final versions of the same consonant are the same. Also, the strict separation does not permit violations of the syllable position constraint that, on the contrary, do occur and appear to be lawful.

5. Null elements and the need for network control of phonological frames. One of the difficulties in the large simulation was that it tended to delete consonants (replace them with a null element) rather than add them, contrary to what is found in error corpora. This is a serious problem because the intrusiveness of the null elements stems from their high syllabic frequency, and the intrusiveness of frequently occurring sounds and sound combinations is, I feel, a desirable aspect of the model. Either the treatment of such frequency effects, which is a basic property of the model, or the null elements have to go.

With this in mind I will briefly sketch a more complete view of the relation between the nodes in the lexical network and phonological frames that does not require null elements and has several advantages over the simulation's treatment. This view is related to Stemberger's (1984c, 1985b) and Shattuck-Hufnagel's (1979) ideas on the process of building phonological frames.

The basic idea is that all form-related lexical nodes (from word to feature) would guide the construction of the appropriate frame for a particular morpheme or word. That is, the decision about how many syllables to create in the frame, what stress value to assign them, and what categorical slots are present in each would be made by addressing the activation levels of certain nodes. In the simulation, the nodes in the network did not participate in frame building. The categories onset, nucleus, and coda were generated over and over again, as long as there was something to say. As I mentioned before, this was a simplification ignoring dependencies between syllabic and metrical structure, effects of word and morpheme boundaries, and much more. However, there was a second problem with the simulation's treatment and that is that many units in the lexical network were largely superfluous. There was no motivation for, say, feature or rime nodes other than to explain feature and rime effects on slips. The more complete view would make use of these nodes' activation levels. So, for example, if the vocalic feature *-tense* is highly activated and the syllable being constructed is word-final, the activated feature would bias for the frame-building process to create a coda slot, because vowels in word-final position must be tense. Rime nodes could, as well, inform the frame-building process in a useful way regarding stress, because a syllable's stress value is, to some extent, predictable from that syllable's rime (Chomsky & Halle, 1968).

In any event, because a wide variety of information from several levels determines the metrical structure of an utterance (Liberman & Prince, 1977), all of the units in the lexical network could have a role in frame building. If frame building is to be guided by the lexical network, there is no longer a need for null elements. A syllable that does not have an onset (or coda) would not have to send activation to a null element that competes with other onsets (or codas). Rather, it would assure the creation of a frame without an onset (or coda). Addition and deletion slips would then result from the wrong frame being built (Stemberger, 1984c). Although much more theory, both linguistic and processing theory, would have to be worked out before a simulation model expressing this view could be constructed, it is, I feel, the next step toward an adequate theory of phonological encoding within the framework of the production theory presented here.

## Issues in Morphological and Syntactic Encoding

According to the theory presented in this article, representations at all levels are built in much the same way. Generative rules define a sequence of categorically defined slots, which are then filled by items that have been retrieved by spreading activation through a lexical network. Each slot is filled by selecting that item of the appropriate category with the highest level of activation. The resulting representation is an ordered set of selected items. This set was imagined to be a collection of order tags attached to nodes in the network that represent the selected items.

So far, the theory has only been worked out in detail in phonological encoding. I have touched on higher level encoding processes briefly while outlining the theory, but have ignored the major issues for the most part. Although I do not develop explicit models of syntactic and morphological encoding, I discuss three preliminary issues, the hypothesized separation between syntactic and morphological processes in the theory, the question of how rules are guided as they build syntactic and morphological frames, and the interactive nature of lexical processes in the theory. By focusing on these issues I hope to illuminate both the problems and the promise of the theory as it applies to higher level phenomena.

### *Separate Syntactic and Morphological Representations?*

There is considerable debate in linguistics and psycholinguistics concerning morphological knowledge. Is the morphology best considered as a separate component or merely the "lower" part of the syntax? There are good linguistic reasons for separating the morphology (both its inflectional and derivational aspects) from the syntax (see Lapointe, 1980; Selkirk, 1982). However, the fact that morphological rules make reference to syntactic categories blurs the distinction somewhat. The constraints on which prefixes and suffixes attach to which stems must refer to the syntactic class of the stem. For example, the prefix *re-* attaches to verbs and suffix *-ness* attaches to adjectives turning them into nouns.

The blurriness of the syntax–morphology distinction in linguistics is reflected in the psycholinguistic evidence as well, particularly in morphemic speech errors. Some morpheme slips appear to be guided by syntactic considerations, whereas others seem to ignore the syntax. Consider word-stem exchanges in which the interacting stems come from distinct clauses and belong to the same syntactic category. An example of such an error would be *Because you blabbed, I'm not staying* being spoken as *Because you stayed, I'm not blabbing.* Because the exchange of stems *strands* the inflectional affixes, *-ing* and *-ed,* the error involves morphemes, not whole words. Despite the stranding of morphemes there are good reasons to assume that this error belongs to the same linguistic level as whole-word misorderings and substitutions such as *writing a mother to my letter* or *Liszt's second Hungarian restaurant* (Garrett, 1980b; Stemberger, 1982a). Furthermore, this level seems to be one as-

sociated with building a syntactic representation of the utterance to be spoken. Evidence that the error *Because you stayed, I'm not blabbing* is syntactic is provided by three features of the error. First, the interacting stems are both verbs. This suggests that the relevant categories at the error's level are syntactic. Recall also, that whole word errors such as those just cited obey the syntactic categorical constraint. Second, the stranded morphemes (*-ing* and *-ed*) are inflectional rather than derivational. On many linguistic accounts, inflectional morphology (but not derivational morphology) is represented as part of syntactic structure (cf. e.g., Chomsky, 1965, 1981). (I should note, however, that at least one alternative account does not assume that inflectional features are manipulated in the syntax; Lapointe, 1980.) Third, the interacting stems are in different clauses. Garrett (1975) suggested that the involvement of forms from different clauses indicates that the error likely belongs to the functional level, which corresponds to the syntactic level here. So, although this error involves morphemes and not whole words, it is possible to locate it in syntactic encoding processes. In Garrett's (1975) model this would be the functional level, and in the model proposed by Stemberger (1982a) the error would be assigned to lexical selection processes.

As we have just seen, some stem misorderings are syntactic in nature. Other morphemic errors appear to be guided less by the syntax and, I will argue, are properly considered as morphological. Consider the following three stem exchanges:

*thinly sliced* → *slicely thinned* (Stemberger, 1982a),
*windows rolled up* → *rolls windowed up* (Garrett, 1975),
and
*all starters scored* → *all scorers started* (Garrett, 1975).

Superficially, the same thing is going on in these errors as in the error *Because you stayed, I'm not blabbing.* Stems were exchanged, stranding their suffixes. However, in these errors the interacting stems were in adjacent words in the same clause. What is more, the interacting stems in the first two errors differ in syntactic class (adjective *thin* vs. verb *slice,* and noun *window* vs. verb *roll*). Although the stems in the third error are both verbs (*start* and *score*), the error differs from the *stayed . . . blabbing* error in that one of the stranded affixes is derivational, agentive *-er.* The *-ly* affix in the first error is also derivational. These and other differences led Garrett (1975) to assign stem exchanges involving close words in the same clause, which often violate syntactic category constraints, to the lower positional level in his theory. This contrasts with exchanges of words and stems from distinct clauses that preserve syntactic class, which were assigned to the functional level. Within the framework of the production theory in this article, the functional stem exchanges are assumed to occur during syntactic encoding and positional exchanges during morphological encoding.

Other morpheme errors also frequently violate syntactic category constraints and thus provide further evidence for a morphological level that can be distinguished from a syntactic level. Suffix-shift errors, such as the two that follow, often involve the movement of the suffix between words of differing syntactic categories (Garrett, 1975):

*he gets it done* → *he get its done* (Garrett, 1975) and
*looking at* → *look atting* (Stemberger, 1982a).

Garrett (1975) argued that these shifts are morphological rather than phonological because shifts of the final parts of words that do not correspond to morphemes are rare. Thus, these errors likely occur at a level that manipulates morphemes, but one that does not seem to be strongly guided by syntactic considerations. For the present theory I will assume that these shifts, along with other affix errors (noncontextual substitutions, anticipations, perseverations, and exchanges) and stem misorderings that violate syntactic constraints occur during the construction of a morphological representation. These errors contrast with errors that are assumed to occur during syntactic encoding (word exchange, anticipation, perseveration, substitution, blend), all of which strongly adhere to the syntactic category constraint. This morphological–syntactic distinction corresponds closely to Garrett's positional–functional distinction. The difference is that I am further distinguishing between the morphological and phonological levels, which together compose Garrett's positional level.

If the construction of the morphological representation is to proceed as specified by the theory, there must be morpheme nodes in the lexical network, one for each morpheme. These must be labeled for their morphological category. The morphological rules should generate frames with slots labeled for morphological category, and each slot should be filled by the morpheme of the appropriate category with the highest activation level.

At this point two questions come to mind. Which units are morphemes? And what are their morphological categories? The first question has been the subject of much psycholinguistic research in both comprehension and production (Butterworth, 1983). Researchers who have focused on morphemic speech errors in English have, in general, concluded that inflectional suffixes, components of compounds, and at least some prefixes and derivational suffixes are genuine morphemic units (e.g., Fromkin, 1971; MacKay, 1979; Stemberger, 1982a). That is, all inflected words (*boys, singing, brought*) and many derived words (*genuineness, clearly*) are in some sense assembled out of morphemes during production. I assume for the theory that at the morphological level, words such as *singing, boys, rewrite, clearly, streetlight,* and *genuineness* are composed of two morphemes. This is supported by the fact that the stems of words such as these can move, stranding their affixes, and that the affixes can interact with other affixes independently of their stems.

Given that many words seem to be polymorphemic at the morphological level, it becomes important to say something about how such words are represented at the syntactic level. Here I will tentatively adopt a linguistic treatment that includes inflectional markers at the syntactic level, but not other affixes (e.g., Chomsky, 1965, 1981). Translating this treatment to the production theory involves the use of nodes for inflections that participate in the syntactic representation. Thus, in the network presented earlier in Figure 1 there is a node for the plural in *swimmers.* Notice also that there is a word node for *swimmer* labeled as a noun, but no single node for *swimmers.* In general, derived, but not inflected, forms exist as word nodes in the lexical network. However, as I mentioned, the derived forms break down into their morphemes at the morphological level. Thus, although *swimmer* is a single word at the syntactic level, it connects to *swim* and agentive *-er* at the morphological level. This

treatment is consistent with the earlier assignment of errors such as *Because you stayed, I'm not blabbing* to the syntactic level. The verb slot in the syntactic representation was filled with the verb *stay* rather than the verb *blab,* and then later because of postselection negative feedback on *stay, blab* is able to be selected in the second clause completing the exchange. The key point is that the inflections are stranded because they are separate nodes from the verbs.

The second question that was raised concerned morphological categories. Morphological rules employ both purely morphological categories (stem, prefix, derivational suffix, etc.) and syntactic categories. Thus, the appropriate categories might be syntactically marked morphological types such as noun stem (*dog, girl*), verb prefix (*re-*), nominalizing verb suffix (*-ment*), inflectional verb suffix (*-ing, -ed*), and so on. There is a problem, though, in adopting such specific categories for the production theory. As we have seen, morphemic errors often violate them. Consider the three errors that follow:

*windows rolled up* → *rolls windowed up,*
*dollars deductible* → *dedollars deductible* (Stemberger, 1982a), and
*some weeks* → *somes weeks* (Stemberger, 1982a).

In the first error, which has already been discussed, noun and verb stems have exchanged. In the second, which could be analyzed as a prefix anticipatory addition, a verb prefix has attached itself to a noun. Similarly, in the third error a noun inflection has been added to a quantifier. The problem with the categories, though, is their syntactic specification, not with their being labeled as stem, prefix, or suffix. In the preceding examples, and generally as well, prefixes move to prefix positions, suffixes to suffix positions, and stems to stem positions (MacKay, 1979). Thus, the category constraint in morphemic errors is similar to that in sound errors, in which onsets, vowels, and codas move only to onset, vowel, and coda positions, respectively. This implies, within the context of the production theory, that the categorically defined slots of the morphological frame are positional, not syntactic.

Thus, there is a dilemma. Linguistically speaking, morphological rules must refer to syntactic categories if they are to generate morphologically acceptable words. But the speech-error data indicate that morphemic errors do not adhere slavishly to the syntactic aspects of morphological categories, only to the positional or structural aspects. A resolution of this dilemma must await more data on the syntactic influences on morphemic errors. There is some indication of a probabilistic tendency for stem misorderings to involve stems of the same syntactic class and for affixes to move preferentially to stems of the appropriate syntactic category (MacKay, 1979; Stemberger, 1982a). From this, one can speculate that the relevant categories code both positional and syntactic information but that the insertion rules that fill the categorically defined slots sometimes ignore the syntactic aspects of the category. These aspects would really only be needed when the speaker is using the morphology productively; that is, when he or she is building a novel word. Otherwise, there would not be a strong need for the insertion process to use syntactic information when selecting and ordering morphemes.

### *How Do Generative Rules Operate?*

The second major issue that must be considered when applying the theory to higher level processes is that of how the generative rules are guided during the building of each representation's frame. This problem has been discussed in the context of activation-based production models for both morphological and syntactic processes (Bock, 1982; Kempen, 1978; Lapointe, 1985; MacKay, 1982; Reich, 1970; Stemberger, 1982a).

At the syntactic level it appears that both input from the semantic representation and input from activated word nodes guide the building of the syntactic frame. Although it is clear that semantic and pragmatic considerations should guide the syntax, it is less clear that the activation levels of word nodes should, as well. There are, however, some experimental findings that indicate that the retrievability of a particular word affects the syntactic structure of the spoken sentence (see Bock, 1982, for review). To model these effects in the present theory, the activation levels of word nodes will have to be taken into account by the syntax. For example, if an intransitive verb is very highly activated, it will have to assure the creation of a frame without a direct object (see Bock, 1982, for a detailed discussion of the interaction between the lexicon and the syntax in production).

By what mechanism, then, are syntactic rules sensitive to lexical activation? One possibility is that the rules themselves form a network that is connected to the lexical network. Activated lexical nodes can then serve as inputs that, along with other inputs, determine when and if particular rules operate (e.g., MacKay, 1982; Reich, 1970). Whether generative rules can be cast in such a network form is an open question (e.g., it is not clear how recursion would be handled), but at the very least I would argue that it is useful to treat syntactic and morphological frame building decisions as resulting from the activation of rule-governed options.

Stemberger (1982b) discussed how some syntactic errors can be seen as resulting from wrongly activated syntactic rules. For example, the word-shift error *roll up it* for *roll it up* could have come from the overly high activation of a syntactic rule that places particles right after their verbs. Thus, the error may have occurred in the frame-building process rather than the process of filling slots in the frame. Although the production theory presented here has focused on slot-filling errors, it is quite compatible with the existence of errors in frame building if frame-building processes are tied to spreading activation.

As an example of how frame building can be tied to spreading activation in the theory, I will present a simple account of how suffix-shift errors can arise from building the wrong morphological frame. Suffix shifts such as *get its* being spoken as *gets it* have been discussed by Garrett (1975) and Stemberger (1982a). These errors tend to involve inflectional, rather than derivational suffixes; the moving suffix nearly always stays in the same clause, and, like other morphological errors, these errors obey a positional categorical constraint—suffixes only shift to suffix positions. For a suffix to be included in a morphological representation, a slot must be created for it. In particular, a morphotactic rule like WORD → STEM SUFFIX must operate as opposed to WORD → STEM. I assume that the latter rule operates unless sufficient activation from morphemic suffix nodes spreads to the morphological rules, in which case a frame with a suffix

slot is built. Figure 10 shows the network configuration for the inflected noun *cats* and for the derived modifier *catty.* In each case, activation has spread from the suffix to a morphotactic rule in charge of building a word frame with a suffix slot. Shifts can thus result if the frame with the suffix is built at the wrong time, when some stem other than *cat* is due to be selected for the stem slot. (Note that additions *some cats* → *somes cats* and deletions *some cats* → *some cat* could also result from the wrong frame being built.) The tendency for shifts to involve inflectional rather than derivational suffixes emerges from the assumption that inflections are not as closely tied to their bases as are derivational suffixes. When a word such as *catty* is morphologically encoded, the morphemes *cat* and *y* are going to be activated together because of the word node *catty,* and thus, the activated *y* node will guide the rules into building a suffix slot just when *cat* is the most highly activated stem. In the case of *cats,* although the plural marker in the syntax will be activated at about the same time as the word node for *cat,* there is a greater chance that the plural morpheme will be activated at a time when *cat* is not the most active stem. Thus, a frame with a suffix slot may be built at the wrong time, shifting the suffix. The fact that suffixes shift to suffix positions within the same clause is accounted for by the categorical specification of suffixes together with the assumption that the syntactic representation is built clause by clause.

This example shows that the spreading-activation assumptions of the theory can be extended beyond the lexical network to explain how frame building is guided and sometimes misguided. The account is sketchy and limited, but it gives a feel for the interaction between rules and activated lexical nodes that a more complete model would call on to explain frame-building errors.

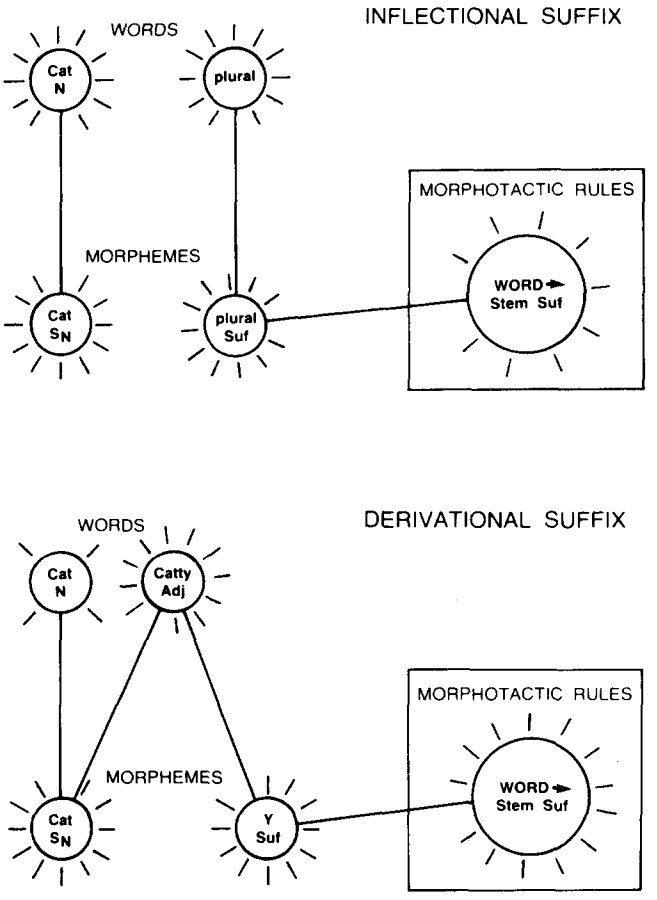


*Figure 10.* Proposed network configuration for inflectional and derivational suffixes. (Activation on the nodes [indicated by highlighting] is consistent with the construction *cats* in the top part of the figure and *catty* in the bottom. All connections are two way.)

## *Interactive Nature of Lexical Processing*

Because of the bottom-up as well as the top-down connections in the network, lexical retrieval is highly interactive. Nodes that participate primarily in later levels of representation can, nonetheless, influence decisions made in earlier levels via bottom-up feedback. The co-occurrence of top-down and bottom-up processes in the lexicon has often been discussed in the context of word recognition (e.g., Tanenhaus & Lucas, in press), and this section discusses how these processes influence lexical selection in the theory.

Consider the syntactic encoding of the sentence *I'm writing a letter to my mother.* Let us assume that the syntactic frame has been constructed and word nodes have been selected for everything before *letter.* At this point the syntax is seeking to fill the noun slot for the direct object of the sentence. The conceptual nodes in the semantic representation that stand for the concept of *letter* should be highly activated. Following Bock (1982), we can assume that they became activated as a result of some signaling between the syntax and the thematic roles that concepts were bound to in the semantic representation and that these concepts have been sending activation to word nodes connected to them. Normally, the word *letter* will be the most activated noun and it will be selected. Other nouns will be active as well, though, and thus are potential competitors for the slot. These fall into four broad classes: (a) nouns with some semantic or pragmatic relation with *letter*—partial synonyms (*note*) or other relations (*stamp, envelope*), (b) nouns from other parts of the intended utterance (*mother*) and words related to them (*father*), (c) nouns representing the extralinguistic environment (If a dog walks into the room, *dog* might be active), and (d) nouns that are phonologically related to *letter* or other highly activated words (*lettuce, litter*).

Why are all of these words active? Those from the first two classes are active simply because their concepts are part of the semantic representation of the utterance, and so their activation comes about from downward spreading. The case is similar for words related to the environment. The relevant concepts are active to the extent that the speaker is attending to these parts of the environment. Phonologically similar words are active because of feedback from common phonological nodes. *Letter* activates the syllable *let* and its phonemes, and their activation feeds back to words and morphemes sharing many of the sounds of *letter.*

Each of these competing nouns could conceivably get more activation than *letter* and replace it. If a word such as *note* replaces it, then there is no identifiable error at all. If *stamp, dog, father,* or *lettuce* replaces it, then a word-substitution error occurs. Such an error could be subcategorized on the basis of the relation between the target and the substituting word. Hence,

the error could be semantic, environmental, or phonological in nature. If another word from the intended utterance, such as *mother,* replaces *letter,* the error is a misordering—in this case either a simple anticipation, *writing a mother to my mother,* or the beginning of an exchange, *writing a mother to my letter.* Thus, word substitutions and misorderings in the theory come from the same mechanism, the filling of slots in the syntactic representation (Dell & Reich, 1981; Stemberger, 1982a, 1985a).

Another fairly common word-level slip is the blend *letter/note → lote.* Following MacKay (1973) and Stemberger (1982a), I assume that it is possible for two word nodes to be nearly equally active and for both to be selected and tagged for the same position. The effect of this on lower encoding levels would likely be a blend of the words' sounds.

In this way the theory can give an informal account of word substitutions, misorderings, and blends, and allows for a number of possible relations between target and substituting words. The real interest in the theory's account of such errors, however, lies in its predictions regarding multiple relations between target and substituting words. Consider the error of saying *Let's stop* for *Let's start.* Is this error semantic or phonological, or is it in some way both? Because of its interactive nature, the theory asserts that the error is likely to be both. Top-down spreading from the conceptual representation for *start* probably provides much of the activation for the word *stop.* However, some bottom-up spreading from the common sounds /st/ also contributes activation to *stop.* One can imagine other sources of activation for *stop,* as well. Perhaps the speaker is torn between saying *Let's start* and *Let's not stop now.* These competing plans could be an important top-down source for *stop* (Baars, 1980; Fay, 1981). In general, the theory emphasizes combinations of sources of activation and hence is compatible with the claim that slips have many causes, as in Freud's principle of overdetermination. More specifically, from the theory we expect word substitutions, blends, and misorderings such as exchanges, anticipations, and perseverations to exhibit more than one relation between the interacting words fairly often. For example, a word substitution might exhibit both an environmental and a semantic effect (*could you get me the spoons, I mean knives?* where the speaker is holding a spoon), or a blend could show both a semantic and a phonological effect (*slick/slippery → slickery*), or an exchange could show a phonological effect (*he eats yoga and does yogurt*). (Note that in misorderings such as exchanges, the interacting words necessarily enjoy a relation of contiguity; that is, they are activated at the same time by virtue of occurring in the same utterance. Hence, the phonological similarity between the words is an additional relation.)

The theory's stress on multiple relations places it in conflict with an important feature of the standard view of lexical selection, which I have called the *two-step theory* of lexical selection (Dell & Reich, 1981). In this theory, which has been advocated by Fromkin (1971) and Fay and Cutler (1977), the main lexicon is ordered such that similar sounding words of the same syntactic category are near to each other. Associated with each word is an address specifying where the main entry of each word is to be found in the listing. So, the addresses of phonologically similar words such as *letter* and *lettuce* would be very close. The first step in selection consists of going from a semantic specification to an address, and the second, of going from the address to the word's phonological form. Semantic word substitutions (*letter → envelope*) happen in the first step, and phonological word substitutions (*letter → lettuce*), in the second step.

Thus, according to the two-step theory, a given error is caused either in the mapping from meaning to address or in the mapping from address to word, but not in both. Essentially, this means that semantic word substitutions and phonological word substitutions should be disjoint sets. Unlike the theory developed in this article, the two-step theory of selection claims that semantic and phonological factors cannot act together to create a word substitution. The present theory, which allows for feedback from sound to word nodes, claims that phonological relations should sometimes be present in clearly semantic errors. This is a specific instance of the theory's general claim of multiple relations in such errors.

The issue of multiple relations in word slips has only recently been investigated, so there is not a great deal of data. What there is, though, supports the existence of multiple relations. Butterworth (1981), Harley (1984), and Dell and Reich (1981) have all found that the interacting words in semantic word-substitution errors are similar in sound to a degree greater than chance. The same was found to be true with blends. Word exchanges and anticipations were also found to exhibit a phonological relation between the interacting words, although the phonological relation in these errors is much weaker than that in word substitutions (Dell & Reich, 1981). It can be concluded that the data are, to a first approximation, consistent with a view of lexical selection such as that contained in the present theory. Furthermore, the existence of phonological effects in semantic errors is inconsistent with noninteractive serial models such as the two-step theory of lexical selection.

## Conclusions: Production, Productivity, and the Causes of Speech Errors

There are several answers to the question of what causes slips of the tongue. An error such as *barn door* being spoken as *darn bore* can be explained as an exchange of the initial consonants /b/ and /d/. This explanation, however, describes the form of the error, not its causes. Within the context of the theory, one can identify several kinds of causes of slips. The main focus of the theory has been on linguistic causes such as the fact that, in the error above, /b/ and /d/ are similar consonants, or that both words have an /r/ in them (the repeated-phoneme effect), or that *darn* and *bore* are both lexical items. Causes such as these do not exhaust the potential causes of error, however. Perhaps the speaker believed that the listener was, in fact, a darn bore. The existence of such Freudian influences on errors has received some empirical support (Motley & Baars, 1979). In addition, one can identify causes such as nervousness, tiredness, or brain damage that have negative influences on a wide spectrum of behavior. (Although such influences are outside of sentence production, one can still use the production theory to understand these effects. For example, does tiredness slow down everything or does it make everything more noisy? Slowing down processing in the theory, by decreasing the spreading parameter, $p$, has different effects on error patterns than does adding noise to the network.)

Each of the factors listed in the preceding paragraph can be thought of as being among the causes of slips. In conclusion, I

will speculate on yet another cause of speech errors, one that stems from the nature of language itself.

Why is the language-production system error-prone? The main reason, I feel, is that the system must be productive. That is, it must allow for the production of novel combinations of items. We recognize that some strings of sounds, although not words, are potential words because they follow phonological rules. The productivity associated with syntactic and morphological rules is even more compelling. This productivity is represented in the theory through the inclusion of generative rules that operate whenever something is said. Note that such rules would not be necessary for a talking machine that possesses a limited set of utterances. Such a machine just needs a set of articulatory instructions for each utterance. The human language-production system is more flexible. But one of the costs of this flexibility is slips of the tongue. In the theory, the locus of a slip is that point where prestored knowledge in the lexical network comes in contact with the productive knowledge represented by the generative rules. This is where the real decisions are made, and hence where the errors are made.

It is interesting to note that where this flexibility is no longer required, we rarely get slips. Consider the productivity of a language's sound system. We seem to have the ability to create and recognize novel combinations of phonemes (it happens as we learn new vocabulary), but we do not have much of an ability to create novel phonemes out of existing features. Another way to say this is that the set of possible phoneme combinations for a language user is open, but the set of possible phonemes is closed. These facts correlate nicely with the commonness of phoneme errors and the rarity of feature errors. One can speculate that the language-production system is organized to allow great flexibility in the ordering and combining of constituents of members of open sets, but not with members of closed sets, and this flexibility comes out in the potential for slips involving the constituents of items from an open set.

## Summary

The theory in this article is an attempt to develop an explicit account of retrieval processes in production, with an emphasis on the retrieval of the phonological form of words. Its strong points include its ability to account for a great deal of existing speech-error data and to lead to quantitative predictions regarding error patterns. These predictions emerge directly from the structure of the theory—from interaction of its linguistic and processing assumptions and not from specific parameters tied to specific effects. Speech errors in the theory are seen not so much as malfunctions, but rather as the consequence of the need for language production to be productive, coupled with some reasonable assumptions about the way that linguistic knowledge is represented and retrieved.

The theory suffers to some extent from a trade-off between explicitness and a desire to avoid unmotivated assumptions. In its account of phonological encoding, the theory is explicit but contains several ad hoc assumptions. Where these assumptions are less in evidence, in the description of higher level processes, the theory becomes a lot fuzzier. This is just another way of saying that much more research needs to be done.

## References


Allen, J. F. (1979). *A plan-based approach to speech act recognition.* Unpublished doctoral dissertation, University of Toronto.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Baars, B. J. (1980). The competing plans hypothesis: An heuristic viewpoint on the causes of errors in speech. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 39–49). The Hague: Mouton.

Baars, B. J., & Motley, M. T. (1974, Fall). Spoonerisms: Experimental elicitation of human speech errors. *Catalog of Selected Documents in Psychology,* Journal Supplement Abstract Service.

Baars, B. J., & Motley, M. T. (1982). *A sympathetic critique of naturalistic studies of slips of speech and action.* Unpublished manuscript.

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status from artifically elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14,* 382–391.

Berg, T. (1983). *Monitoring via feedback in language production: Evidence from cut-offs.* Unpublished manuscript.

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

Bock, J. K. (1983, November). *Priming syntactic form in sentence production.* Paper presented at the meeting of the Psychonomic Society, San Diego, CA.

Boomer, D. S., & Laver, J. D. M. (1968). Slips of the tongue. *British Journal of Disorders of Communication, 3,* 1–12.

Brown, R., & McNeill, D. (1966). The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Butterworth, B. (1981). Speech errors: Old data in search of new theories. *Linguistics, 19,* 627–662.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257–294). London: Academic Press.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: M.I.T. Press.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures.* Dordrecht, Netherlands: Foris.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper.

Cole, R. A., Jakimik, J., & Cooper, W. E. (1978). Perceptibility of phonetic features in fluent speech. *Journal of the Acoustical Society of America, 64,* 44–56.

Cooper, W. E., Egido, C., & Paccia, J. M. (1978). Grammatical control of a phonological rule: Palatalization. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 264–272.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

Cottrell, G. W., & Small, S. L. (1983). A connectionist scheme for modelling word sense disambiguation. *Cognition and Brain Theory, 6,* 89–120.

Crompton, A. (1981). Syllables and segments in speech production. *Linguistics, 19,* 663–716.

Cutler, A. (1980). Errors of stress and intonation. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 67–80). New York: Academic Press.

Cutler, A. (1981). The reliability of speech error data. *Linguistics, 19,* 561–582.

Dell, G. S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 222–233.

Dell, G. S., & Reich, P. A. (1977). A model of slips of the tongue. *The Third LACUS Forum, 3,* 448–455.

Dell, G. S., & Reich, P. A. (1980). Toward a unified theory of slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 273–286). New York: Academic Press.

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.

Dosher, B. A. (1982). Effect of sentence size and network distance on retrieval speed. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 173–207.

Fay, D. (1981). Substitutions and splices: A study of sentence blends. *Linguistics, 19,* 717–750.

Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8,* 505–520.

Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science, 6,* 205–254.

Ford, M. (1978). *Planning units and syntax in sentence production.* Unpublished doctoral dissertation, University of Melbourne, Australia.

Ford, M., & Holmes, V. (1978). Planning units and syntax in sentence production. *Cognition, 6,* 35–53.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 373–420). London: Academic Press.

Freud, S. (1958). *Psychopathology of everyday life* (A. A. Brill, Trans.). New York: New American Library. (Original work published 1901)

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

Fromkin, V. A. (1973). Introduction. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 11–45). The Hague: Mouton.

Fromkin, V. A. (Ed.). (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand.* New York: Academic Press.

Fry, D. B. (1969). The linguistic evidence of speech errors. *Brno Studies in English, 8,* 69–74.

Garnham, A., Shillcock, R. C., Brown, G. D. A., Mill, A. I. D., & Cutler, A. (1981). Slips of the tongue in the London–Lund corpus of spontaneous conversation. *Linguistics, 19,* 805–817.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales, & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–255). Amsterdam: North-Holland.

Garrett, M. F. (1980a). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 177–220). London: Academic Press.

Garrett, M. F. (1980b). The limits of accommodation: Arguments for independent processing levels in sentence production. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 263–271). New York: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Gazdar, G. (1980). Pragmatic constraints on linguistic production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 49–68). London: Academic Press.

Goldstein, L. (1977). *Features, salience, and bias* (UCLA Working Papers in Phonetics No. 39). Los Angeles: University of California.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology, 11,* 58–81.

Grossberg, S. (1982). Studies of mind and brain: *Neural principles of learning, perception, development, cognition, and motor control.* Boston: Reidel.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93,* 46–74.

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

Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 97–110). New York: Academic Press.

Kempen, G. (1978). Sentence construction by a psychologically plausible formulator. In R. N. Campbell, & P. T. Smith (Eds.), *Recent advances in the psychology of language.* New York: Plenum Press.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185–209.

Kupin, J. (1979). *Tongue twisters as a source of information about speech production.* Unpublished doctoral dissertation, University of Connecticut.

Lamb, S. (1966). *An outline of stratificational grammar.* Washington, DC: Georgetown University Press.

Lapointe, S. (1980). *A theory of grammatical agreement.* Unpublished doctoral dissertation, University of Massachusetts, Amherst.

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

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. (1982). Linearization in describing spatial networks. In S. Peters & E. Saarinen (Eds.), *Processes, beliefs, and questions* (pp. 199–220). Dordrecht, Netherlands: Reidel.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41–104.

Levelt, W. J. M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. Klein & W. J. M. Levelt (Eds.), *Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch* (pp. 221–252). Dordrecht, Netherlands: Reidel.

Levitt, A. G., & Healy, A. F. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language, 24,* 717–733.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8,* 249–336.

Lockwood, D. G. (1972). *Introduction to stratificational linguistics.* New York: Harcourt, Brace, Jovanovich.

MacKay, D. G. (1969). Forward and backward masking in motor systems. *Kybernetik, 2,* 57–64.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8,* 323–350.

MacKay, D. G. (1971). Stress pre-entry in motor systems. *American Journal of Psychology, 84,* 35–51.

MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology, 3,* 210–227.

MacKay, D. G. (1973). Complexity in output systems: Evidence from behavioral hybrids. *American Journal of Psychology, 86,* 785–806.

MacKay, D. G. (1979). Lexical insertion, inflection, and derivation: Creative processes in word production. *Journal of Psycholinguistic Research, 8,* 477–498.

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

Marslen-Wilson, W., & Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.

Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.

McClelland, J. L. (1979). On the time relations of mental processes: An

examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

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

Meringer, R., & Mayer, K. (1895). *Verspechen und Verlesen.* Stuttgart, FRG: Goschensche.

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.

Motley, M. T., & Baars, B. J. (1975). Encoding sensitivities to phonological markedness and transition probability: Evidence from spoonerisms. *Human Communication Research, 2,* 351–361.

Motley, M. T., & Baars, B. J. (1976). Semantic bias effects on the outcomes of verbal slips. *Cognition, 4,* 177–188.

Motley, M. T., & Baars, B. J. (1979). Effects of cognitive set upon laboratory induced verbal (Freudian) slips. *Journal of Speech and Hearing Research, 22,* 421–432.

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 21,* 578–594.

Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. Van Raad (Eds.), *Leyden studies in linguistics and phonetics* (pp. 114–132). The Hague: Mouton.

Nooteboom, S. G., & Cohen, A. (1975). Anticipation in speech production and its implications for perception. In A. Cohen & S. Nooteboom (Eds.), *Structure and process in speech perception,* New York: Springer.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88,* 1–15.

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review, 88,* 454–462.

Reich, P. A. (1970). The English auxiliaries: A relational network description. *Canadian Journal of Linguistics, 16,* 18–50.

Reich, P. A. (1977). Evidence for a stratal boundary from slips of the tongue. *Forum Linguisticum, 2,* 211–217.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274–280.

Roche-Lecours, A., & L'hermitte, F. (1969). Phonemic paraphasias: Linguistic structures and tentative hypotheses. *Cortex, 5,* 193–228.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive motor performance. *Cognitive Science, 6,* 1–36.

Selkirk, E. O. (1982). *The syntax of words.* Cambridge, MA: MIT Press.

Shastri, L., & Feldman, J. A. (1984). *Semantic networks and neural nets* (Tech. Rep. No. 131). Rochester, NY: University of Rochester, Computer Science Department.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Erlbaum.

Shattuck-Hufnagel, S. (1981, December). *Position of errors in tongue twisters and spontaneous speech: Evidence for two processing mechanisms?* Paper presented at 102nd meeting of the Acoustical Society of America, Miami.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109–136). New York: Springer-Verlag.

Shattuck-Hufnagel, S., & Klatt, D. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18,* 41–55.

Stemberger, J. P. (1982a). *The lexicon in a model of language production.* Unpublished doctoral dissertation, University of California, San Diego.

Stemberger, J. P. (1982b). The nature of segments in the lexicon: Evidence from speech errors. *Lingua, 56,* 235–259.

Stemberger, J. P. (1982c). Syntactic errors in speech. *Journal of Psycholinguistic Research, 11,* 313–345.

Stemberger, J. P. (1983). *Speech errors and theoretical phonology: A review.* Bloomington, IN: Indiana University, Linguistics Club.

Stemberger, J. P. (1984a). *Lexical bias in errors in language production: Interactive components, editors, and perceptual biases.* Unpublished manuscript, Carnegie-Mellon University, Pittsburgh.

Stemberger, J. P. (1984b). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology, 1,* 281–313.

Stemberger, J. P. (1984c). *Wordshape errors in language production* (Research on Speech Perception Progress Rep. No. 10, pp. 265–300). Bloomington, IN: Indiana University, Speech Research Laboratory.

Stemberger, J. P. (1985a). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–186). London: Erlbaum.

Stemberger, J. P. (1985b). *Phonological rule ordering in a model of language production.* Bloomington, IN: Indiana University, Linguistics Club.

Stemberger, J. P., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language, 25,* 163–180.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typing. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.

Tanenhaus, M. K., & Lucas, M. (in press). Context effects in word recognition. *Cognition.*

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15,* 49–74.

Treiman, R. (1984). On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior, 23,* 343–356.

Wells, R. (1951). Predicting slips of the tongue. *Yale Scientific Magazine, 3,* 9–30.

Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review, 76,* 1–15.

Received May 10, 1984
Revision received September 5, 1985 ■