3
The Sequencing of Action

> Not only speech, but all skilled acts seem to involve the same problems of serial ordering, even down to the temporal coordination of muscular movements in such a movement as reaching and grasping. Analysis of the nervous mechanisms underlying order in the more primitive acts may contribute ultimately to the solution of even the physiology of logic. . . . Serial order is typical of the problems raised by cerebral activity; few, if any, of the problems are simpler or promise easier solution. We can, perhaps, postpone the fatal day when we must face them, by saying that they are too complex for present analysis, but there is a danger here of constructing a false picture of those processes we believe to be simpler.
>
> (Lashley, 1951, pp. 122, 197)

Any theory of action must deal with three basic questions: What is the structure of the components representing skilled behavior? How are these components activated in proper sequence? And how are these components timed or produced at the appropriate rate? The previous chapter examined the first of these problems, the structure of the components for organizing everyday actions, and the present chapter examines the second, how these components become activated in proper sequence.

I begin by outlining the general requirements for a theory of sequencing in action. I then develop a theory that meets these general requirements and makes new predictions for future test. In Chapter 4, which deals with perceptual processes, I examine the related problem of sequencing in perception.

The Sequencing of Action

How do we execute sequences of behavior in proper serial order when we do and in improper order when we make errors? As Lashley (1951) pointed out, sequencing is a general problem for psychological theories. Any behavior more complex than a spinal reflex is sequentially organized and requires explanation in a general theory of sequencing. However, speech provides the most

The sequencing of action. Ch 3 (pp. 39-61) in MacKay, D.G. (1987). ***The organization of perception and action: A theory for language and other cognitive skills*** **(1-254). Berlin: Springer-Verlag.**

extensively studied example of the sequencing problem. Other cerebral activities may employ similar sequencing mechanisms (Lashley, 1951; D. G. MacKay, 1985), but sequencing is especially complex and interesting in the case of speech production because sequencing issues arise at many different levels at once. In discourse, how do we produce sentences one after the other in logical order? How do we order the words within the sentences? How do we order the morphemes, syllables, and segments that make up the words? Finally, how do we order the muscle movements that give rise to the sequence of sounds? Our everyday capacity to organize and produce such a hierarchy of simultaneous, nested sequences is probably fundamental to our uniquely human ability to use language (see also Keele, 1987).

Besides being a multilevel and omnipresent issue, sequence also plays an essential role in all languages. How the components at any given level are sequenced makes a fundamental difference to the significance of an utterance. Whatever the language, changing the order of phonemes in a word, for example, changes the meaning of the word. Neither the omnipresence nor the significance of sequence is unique to language, however; both characteristics are apparent in many other everyday activities. Consider the normal sequence of actions required to light a candle, for example: (1) light the match, (2) apply the match to the wick of the candle, and (3) blow out the match. Performing these actions in any other order, such as (1) light the match, (2) blow out the match, and (3) apply the match to the wick of the candle, is analogous to producing a nonsense word. The overall behavior changes in significance when its component actions are performed in this new order.

General Requirements for a Theory of Sequencing

Any theory of sequencing must address a relatively small number of fundamental questions: Is there a nonsequential or preparatory stage that precedes the sequential activation of behavior? What is the relationship between the sequencing mechanism and the output units for producing behavior? Can sequencing be accomplished by mechanisms responsible for timing? What is the relationship between mechanisms for sequencing and timing in behavior? As we will see, available data bearing on these questions impose general constraints on all viable theories of sequencing.

Preparation for Sequencing

Lashley (1951) was the first to recognize that a priming or preparation stage is necessary for sequencing. According to Lashley, a set of output units must be primed or simultaneously readied for activation before an independently stored sequencing mechanism can activate and impose order on them. Lashley (1951) outlined three sources of support for his idea that simultaneous priming precedes sequential activation. One was anticipatory errors in speech, where an upcoming or soon-to-be-produced word or speech sound becomes produced before its time. An example is, "We have a laboratory, I mean, computer in our own laboratory." Anticipatory errors are the most frequent general class of speech errors and indicate that prior to actual activation soon-to-be-produced units are simultaneously preexcited, primed, or readied for activation. Otherwise, why would an upcoming or about-to-be-produced unit be so much more likely to intrude than any other unit in the speaker's vocabulary?

Another argument for a (simultaneous) preparatory stage prior to (sequential) activation is that "a general facilitation, a rise in the dynamic level" seems necessary for the performance of many sequential activities (Lashley, 1951, p. 187). For example, when sufficiently aroused, brain-damaged patients can execute sequences of behavior that under normal circumstances they cannot. An aphasic who is unable to produce the word "watch" in a laboratory test may exclaim, "Give me my watch!" when someone pretends to make off with his or her watch (H. L. Teuber, personal communication, April, 1965). Such examples suggest that an output sequence cannot become activated unless its units become sufficiently primed. Of course, motivational factors contribute to the required level of priming in this particular example from neuropsychology, whereas in general, priming normally arises mainly from factors specific to the action being produced.

Lashley also noted evidence for priming in studies of reaction time and word association.

> Reaction time, in general, is reduced by preliminary warning or by instructions which allow the subject to prepare for the specific act required. In controlled association experiments, the subject is instructed to respond to the stimulus word by a word having a certain type of relation to it, such as the opposite or a part of which the stimulus is the whole: black–white, apple–seed. The result is an attitude or set which causes the particular category to dominate the associative reaction. (1951, p. 187)

It is as if controlled association instructions simultaneously prime or ready for activation an entire category of specific responses, thereby short-circuiting the first stage of the prime-then-activate process, so that the appropriate response can be produced soon after presentation of the stimulus.

Lashley's (1951) third basis for assuming that priming precedes sequential activation during production is that perception exhibits a similar process. To demonstrate perceptual priming, Lashley auditorily presented to his audience the garden path sentence, "Rapid righting [writing] with his uninjured hand saved from loss the contents of the capsized canoe." As might be expected, a sudden reinterpretation of the word *writing (righting)* took place once the audience heard the last two words of the sentence (see also Carroll, 1986). On the basis of this demonstration, Lashley argued that the units for comprehending the concept "righting" (rather than "writing") could not become activated until the phrase "capsized canoe" had occurred and so must have been held in a state of readiness or partial activation "for at least 3 to 5 seconds after hearing the word" (1951, p. 193). Thus, priming, or readying for activation, precedes actual activation during comprehension, and by analogy during production as well, because "the processes of comprehension and production of speech have too much in common to depend on wholly different mechanisms" (1951, p. 186).

INDEPENDENCE OF SEQUENCE AND CONTENT

The mechanism for sequencing behavior must be separate from the units that represent the content or form of the behavioral sequence; the basic units representing perception and action must be independent of their sequencing mechanism. To see why this is so, consider a class of theories where sequencing and content are nonindependent: chain-association theories. There is no content-independent sequencing mechanism in chain-association theories. Unidirectional links between content nodes provide the representation of sequence. Activating the first content node directly causes activation of the second (connected) content node, and so on, until the entire sequence has been produced.

Many variants of this unidirectional bond assumption have been proposed, and the bonds are usually excitatory in nature. But not always. For example, Estes (1972) proposed a chain-association theory where the bonds are inhibitory rather than excitatory. The first unit inhibits the remaining units, the second inhibits all but the first, the third inhibits all but the first two, and so on. For example, in producing a simple word such as *act*, a superordinate node representing the entire word becomes activated and primes its three subordinate nodes representing the segments, /a/, /c/, and /t/. Now under the unidirectional bond assumption, the first element, representing /a/, inhibits the other two, and the second element, representing /c/, inhibits the third, representing /t/. Thus, the first element, not being inhibited by any of the others, achieves the greatest degree of priming, and becomes activated. The second, no longer being inhibited by the first, now has the greatest priming and becomes activated, releasing the third from inhibition, and so on.

Lashley (1951) pointed out the basic problem with this and related chain-association proposals. The problem is that links between the basic output components will interfere with one another. For example, either excitatory or inhibitory links between the components for *act* will prevent error-free production of *cat* and *tack*, or any other words containing the same components in a different order and vice versa. Extrapolating to a normal 50,000-word vocabulary, conflicting connections between the basic output components would simply prevent speech altogether. Of course, it might be suggested that sequential connections between content nodes are not permanent but are established on the spot as part of the preparation for sequencing (Norman & Rumelhart, 1983). However, this suggestion simply begs the question and adds a new unresolved issue: How are the appropriate (and no other) connections formed, and how are they formed so quickly?

Theories postulating nonindependent sequencing and content units therefore fail to explain the production of sequence per se. These theories also predict errors that do not occur and have difficulty explaining the errors that do occur (D. G. MacKay, 1970e). Because they postulate nonindependent mechanisms for sequence and content, these theories also have difficulty explaining flexibility in sequential behavior. Children's word games, for example, Pig Latin, illustrate the nature of this flexibility (D. G. MacKay, 1972). When playing Pig Latin, children quickly and easily impose a new order on the segments of both never previously encountered nonsense syllables, such as *snark*, and frequently used words, such as *pig* (see also Treiman, 1983). When children produce the word *pig* as *igpay*, for example, no painful process of unlearning the old habitual sequence is required, as might be expected if the old sequence were built into the output units themselves by means of undirectional links. Instead, the sequencing mechanism appears to operate on the basis of internal rules that can be easily altered and that can apply to an indefinitely large number of behavioral units, including never previously encountered ones such as *snark*.

Lashley (1951) noted one final set of phenomena calling for a rulelike sequencing mechanism that is independent of the content units themselves: the ability to translate freely from one language to another using different word orders. An experienced translator does not have to proceed word by word, but can and often must rapidly alter the order of the components making up the original idea. Such flexibility suggests that sequence is not part of the ideas per se, but is imposed on these ideas by a language-specific sequencing mechanism. The way that bilinguals sometimes impose the *wrong* order on words likewise suggests that the sequencing mechanism is independent of the words and ideas being sequenced. A native speaker of German may sometimes impose aspects of German syntax when speaking English, postponing the verb to the end of a familiar English expression, for example. One might argue about whether and how the sequencing instructions for one language can become attached to the word units for the other language in these examples, but such errors simply could not occur if sequencing mechanism and word units were inseparable.

SEQUENCING AND THE INITIATION OF BEHAVIOR

Theories of sequencing must explain a special and repeatedly demonstrated relationship between sequencing and the initiation of behavior. Studies such as Klapp et al. (1973), Sternberg et al. (1978), and Klapp and Wyatt (1976) have shown that it takes less time to initiate a preplanned behavior that consists of a single component than one that consists of a sequence of components. This relationship between sequencing and the initiation of behavior is an embarrassment to chain-association or horizontal link theories, even those augmented with vertical links (e.g., Estes, 1972; Wickelgren, 1979). It also presents problems for theories incorporating a scanning mechanism. In scanning theories, such as D. G. MacKay (1971), a behavioral sequence is loaded into a memory buffer in preparation for sequencing, and behavior becomes initiated by a scanner that sweeps over the buffer from, say, left to right. Thus, a subject who is prepared to say the word *paper*, for example, has already loaded *paper* into the output buffer, and following a go or "speak now" signal, the scanner sweeps over the buffer, causing activation of the initial /p/, followed in turn by the remaining segments of the word. This sequencing process is of course independent of word length, so that the scanner should trigger the initial /p/ of a one-syllable word such as *paint* no faster than the initial /p/ of a two-syllable word such as *paper*.

Available data do not support this prediction, however. As noted in the previous chapter, production onset time is significantly longer for two-syllable words such as *paper* than for one-syllable words such as *paint* (Klapp et al., 1973, among others). Klapp and Wyatt (1976) also observed a similar relationship between sequencing and the initiation behavior in production onset times for sequences of finger movement. Subjects in Klapp and Wyatt (1976) produced one of four Morse code sequences on a telegraph key: *dit-dit*, *dit-dah*, *dah-dit*, and *dah-dah* (to produce a *dit*, the key is released immediately after the press, and to produce a *dah*, the key is held down for about 200 ms prior to release). A light indicated which of the four response sequences to produce, dependent variables being production onset time and the time between the first and second response components. The subjects were of course college students with no prior experience in generating Morse code.

There were three main results: (1) Production onset time for sequences beginning with *dit* was shorter than for sequences beginning with *dah*, but the nature of the second response (*dit* versus *dah*) had no effect on production onset time. (2) The time to initiate the second response (following the first) was longer for *dah* than for *dit*. (3) Production onset time was much faster, however, when the second component was identical to the first (both *dits* or both *dahs*) than when the second was different (e.g., a *dit* and a *dah*).

To explain these results, Klapp and Wyatt (1976) reasoned that planning a *dit* was simpler than planning a *dah*, that only the first response was planned during production onset time, and that the second response was planned during the interresponse interval following the first. However, the third observation contradicted this explanation. Because production onset time was shorter when the second response was the same as the first, the second response must have been planned prior to initiating the first. This seemingly contradictory finding has remained unexplained since Klapp and Wyatt's (1976) study.

ERRORS IN SEQUENCING

Theories of sequencing must of course explain how sequential errors occur—not just the fact that sequential errors occur, but the detailed nature of the regularities that have been observed in these errors. An example is the sequential class phenomenon, one of the strongest and most general regularities observed to date in speech errors. The phenomenon is this: When a speaker inadvertently substitutes one linguistic component for another, both components usually belong to the same sequential class. Cohen (1967) originally observed this regularity in errors involving interchanged words. An example is the error, "We have a laboratory in our own computer," where one noun (*laboratory*) interchanges with another (*computer*). As in this example, nouns generally interchange with other nouns, verbs with verbs and not with, say, nouns or adjectives (Cohen, 1967). Even "Freudian slips" such as, say, "He found her crotch, I mean, watch" (example modified from Fromkin, 1973), adhere to this sequential class rule. Because both *watch* and *crotch* are nouns, this example obeys the sequential class regularity, even though, as Fromkin (1973) points out, semantic (Freudian) factors may also have played a role.

The sequential class regularity has also been observed for errors involving the following components: (1) Morphological components: prefixes interchange with other prefixes, suffixes with other suffixes, and never prefixes with suffixes (D. G. MacKay, 1979). (2) Syllabic components: initial consonant clusters interchange with other initial clusters, and final with final, but never initial with final (D. G. MacKay, 1972; semivowels of course excluded, see Stemberger, 1983). (3) Segmental components: vowels interchange with vowels, consonants with consonants, and never vowels with consonants (D. G. MacKay, 1972). In short, the sequential class regularity holds for all levels of speech production, and a viable theory of sequencing must explain this fact.

Practice Effects

Why do some behaviors exhibit sequencing errors, but not others? For example, humans don't make purely sequential errors in walking and neither do horses. Similarly, we almost never make sequential errors involving the phonological components of function and content words, producing *ce that* for *the cat*, for example. Frequency/practice is almost certainly responsible for this function word effect (see the recent experiments of Dell, 1985b) and may contribute to the absence of sequential errors in walking as well (D. G., MacKay, 1982).

DIFFERENT MECHANISMS FOR SEQUENCING AND TIMING

Another general constraint on theories of sequencing is that different mechanisms are required to time and to sequence behavior. Sequencing cannot be achieved by a timing mechanism, and timing cannot be achieved by a sequencing mechanism. In what follows, I examine both of these hypothetical possibilities in turn to show why neither works.

Consider first the possibility (proposed by Rosenbaum, 1985) that a timing mechanism is by itself responsible for both sequencing and timing in speech production. This hypothetical timing mechanism is able to generate the sequence of phonemes in a word by specifying their time of production, and sequencing errors arise because phonemes have been assigned improper times. For example, the word *cat* might be misproduced as *act* because the *a* has been produced relatively early and the *c* produced relatively late. Likewise, at a higher level, the phrase "in the car," might be misproduced as "in car the," because the noun is produced relatively early, and the article relatively late.

The problem with this account is that in general, substituted components in actually occurring sequential errors don't just exchange places in time; sequential class almost invariably plays a role. For example, in the error "cake the ring of teas" instead of "take the ring of keys," the segments /t/ and /k/ exchange temporal positions, but they also belong to the same domain or sequential class, *initial consonant group*. Even in haplologies or "skipping errors" – such as *shrimp and egg* misproduced as *shrigg*, skipping *–mp and e–* (from Stemberger, 1985) – the speaker skips to a component in the same sequential class (final consonant group in this example) as was required for the intended word. These and other sequential regularities (Stemberger, 1985) would not be expected if a timing mechanism determines sequencing.

Consider now the other hypothetical possibility, that a sequencing mechanism determines both sequencing and timing, an idea proposed by Norman and Rumelhart (1983), among others. Norman and Rumelhart's (1983) theory of typing incorporates a sequencing mechanism, but no timing mechanism, and timing depends on the nature of the operations required for sequencing. Under this view, errors in the timing and sequencing of typestrokes are one and the same. When typestrokes occur out of sequence, one component is being activated especially early, and the other is being activated especially late.

A critical piece of data contradicting this hypothesis appears in Grudin (1981). Grudin had skilled typists type a large corpus of text, and examined their keystroke intervals, the time between one keystroke and the next. He was especially interested in the keystroke intervals for inadvertently produced transposition errors, where *the* is mistyped as *hte*, for example. The results, averaged over a large number of two-letter transpositions, showed no tendency for one key to come especially early and the other especially late. Rather, the keys exchanged places both in sequence, and in time, just as in speech errors. For example, assume that a skilled typist normally types the word *the* correctly with about 140 ms between hitting *space* and *t* and 75 ms between hitting *t* and *h*. Grudin found that when this typist produced the transposition error *hte*, timing remained the same, about 140 ms between space and *h* and 75 ms between *h* and *t*. The wrong components were activated at the right time. This finding indicates that timing is independent of the behavior being timed, and this independence could only occur with separate mechanisms for determining the content, sequencing, and timing of behavior.

Grudin's (1981) findings also indicate that timing is being "programmed" in proficient typing, and this is an especially important fact for theories of sequencing and timing, because typing is a skill that does not demand consistent or accurate timing, unlike, say, music or Morse code. Apparently a timing mechanism plays a role in skilled behavior even when precise and consistent timing is unnecessary.

CLOSE RELATIONSHIP BETWEEN SEQUENCE AND TIMING

A final requirement for theories of sequencing is a close relationship between mechanisms for sequencing and mechanisms for timing; even though timing and sequencing mechanisms are independent of one another, they must nevertheless be closely connected. Findings of Schmidt (1980) can be interpreted as providing preliminary evidence for this close connection. Schmidt had subjects practice moving a lever to a target in a specified period of time, giving them feedback on their movement time following each trial. Different subjects practiced two different types of movement: one was sequential, the other nonsequential. In the nonsequential condition, subjects simply moved the lever horizontally to the specified target in the specified amount of time. In the sequential condition, two movements were required: first to the target and then back to the start position, again in a specified period of time.

The independent variable was a change in the mass of the lever. Suddenly, and without warning, the lever became more difficult to move. This change had strikingly different effects on sequential versus nonsequential movements. In completing the nonsequential movement with the sluggish lever, the subjects reached the target but took longer to do so. However, in completing the sequential movement, they finished at the correct time but undershot the target in space. When producing highly practiced *sequential* movements, we program the time to produce the components of the sequence in advance, and we find it difficult to change this preprogrammed timing, as if sequence and timing involve closely coupled mechanisms.

The Node Structure Theory of Sequencing

In addition to explaining the constraints discussed above, a node structure theory must also explain how nodes with the dynamic properties (e.g., priming, activation, and linkage strength) and the structure of interconnections discussed in previous chapters give rise to the sequential organization of rapidly produced actions. As others have noted, this is a major problem for parallel distributed processing (PDP) theories (see McClelland et al., 1986).

The fact that content nodes are hierarchically organized in the node structure theory means that more nodes must be activated in sequence than if content nodes were organized on only a single level. As we have seen, however, sequencing which is applied at a single level cannot explain natural skills such as speech production. Moreover, the benefits of hierarchical organization greatly outweigh the costs in additional sequence nodes and sequential decisions. Hierarchies facilitate creativity and flexibility of expression; different lower level expressions of the same higher level content become possible in a hierarchic skill. As Keele (1985) and others have pointed out, hierarchies also cut down on how much new learning or connection formation is required. Preformed lower level node structures simply become attached to new higher level nodes and used for new purposes. For example, no new learning is required at the phonological level when adults encounter a new word formed by reassembling morphological components of already familiar words (see also D. G. MacKay, 1982).

The theory I develop here straightforwardly extends the theory I proposed (D. G. MacKay, 1982) for explaining how practice makes behavior more fluent (faster, less prone to error) and more flexible (adapting readily to changed circumstances and transferring readily from one response mechanism to another). Only minor modifications have been necessary for the purpose of developing a unified theory of perception and action.

Sequencing in Action Hierarchies

The content nodes in Figure 3.1 illustrate the problems of sequencing and nonsequencing as they apply to the node structure theory. Some of the nodes in Figure 3.1 must be activated simultaneously rather than sequentially. Distinctive feature nodes receiving simultaneous priming from the same segment node represent an example. All of the distinctive feature nodes for the /p/ in *prove* must

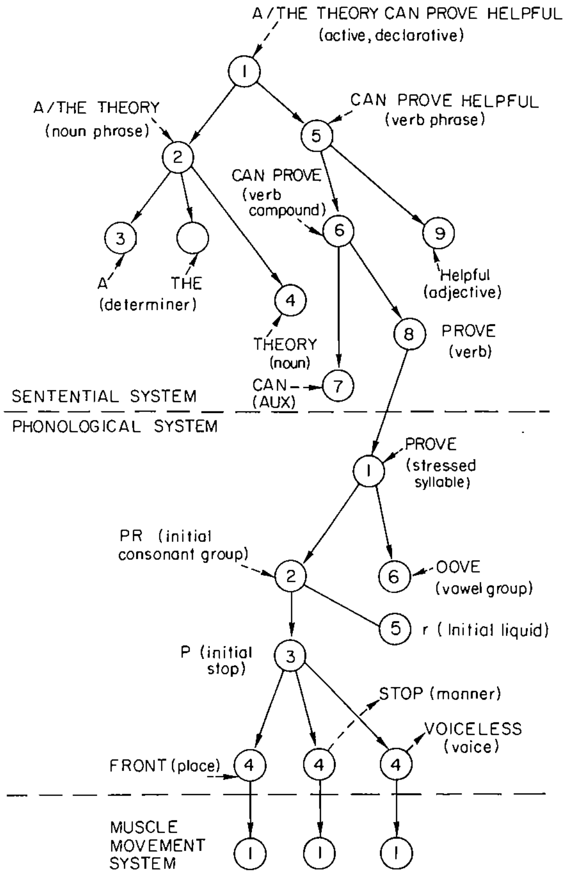

FIGURE 3.1. Aspects of the action hierarchy for producing the sentence "A/the theory can prove helpful." Note that node activation is sequenced within a system, so that the numbers within nodes denote relative order of activation within their respective systems.

become activated at the same or nearly the same time. Achieving this simultaneity constitutes a serial-order problem in reverse.

However, most of the nodes in Figure 3.1 must be activated in sequence if the output is to be error free, and the numbers within the relevant nodes represent the order of activation within a system. The highest level node, *a/the theory can prove helpful*(active declarative), must be activated first. This simultaneously primes both *a/the theory*(noun phrase) and *can prove helpful*(verb phrase), but only *a/the theory*(noun phrase) must become activated at this point. Activating this node simultaneously primes its connected nodes, *a*(determiner), *the*(determiner), and *theory*(noun).

Similarly, activating *prove*(verb) primes a syllable node in the phonological system, which in turn primes *pr*(initial consonant group). *Activating pr*(initial consonant group) primes two connected nodes representing the segments *p* and *r*. Activating *p*(initial consonant) primes a set of phonological feature nodes, including one representing the frontal place of articulation of *p*. Finally, activating the frontal feature node primes a set of muscle movement nodes, including one for contracting the obicularis oral muscles of the lips. Full-fledged behavior of course only takes place if, and only if, the muscle movement nodes at the lowest level of the action hierarchy become activated. The issue, then, is this: How are all these nodes activated in proper order?

The Sequencing Mechanism

So far, the book has focused mainly on the basic components of the theory, content nodes for organizing the form or content of a preplanned action. I have explained why content nodes are hierarchically organized and why they each represent a class of actions, so that the content node representing, say, the phoneme /p/, controls all of the context-dependent ways of pronouncing a /p/, including whispering and shouting. I now elaborate on sequence nodes, their rationale, the structure of their connections, their relationship to content nodes, and the differences between sequence and content nodes in the theory. I then apply the node structure theory of sequencing to specific types of data and outline some predictions of the theory for future test.

Sequence nodes are the triggering mechanisms that determine whether, and in what order, the content nodes in an action hierarchy become activated. The sequencing process involves categories or domains of content nodes, rather than individual content nodes, as in chain-association theories. Recall that a domain consists of all of the nodes connected to a single sequence node. Stated differently, all of the nodes in a domain become activated by means of the same triggering mechanism. The sequence code COLOR ADJECTIVE, for example, is connected to, and serves to activate the many content nodes in the domain *color adjective*. When activated, a sequence node multiplies the priming of every node connected with it by some large factor within a relatively brief period of time. (For discussions of multiplicative gating processes and their possible physiological implementations, see Grossberg, 1982; Sejnowski, 1981.) This multiplicative process has no effect on an unprimed node but soon serves to activate (i.e., bring to threshold) the content node with the greatest degree of priming in its domain, normally the one that has just been primed from above via a connection from a superordinate content node. In producing the adjective *green*, for example, *green*(color adjective) must first become primed, either from above, via a superordinate node such as *green apples*(noun phrase), or from below, via, say, visual perception of either the color green or the printed word *green*. A content node receiving first-order priming passes second-order priming to its connected sequence node, in this case, COLOR ADJECTIVE. Then the sequence node becomes activated (by a process discussed in the next section) and in turn activates the most primed content node in its domain, in this case, *green*(color adjective). This most-primed-wins principle is extremely general and governs the activation of sequence nodes as well as content nodes (D. G. MacKay, 1982).

Without becoming quenched, a sequence node can self-sustain its activation for a set period, say 20 ms, and multiplies the priming of its connected nodes by, say, a factor of 2 every millisecond. The original degree of priming for a content node must therefore exceed some minimal level, so that multiplication of priming by the sequence node over 20 ms can achieve the threshold level required for self-sustained activation of the content node. If the threshold level is 100, this minimal level is of course 100 divided by 2^{20} in this example. Below this minimal level, the multiplied priming of the content node remains subthreshold, and activation cannot occur. And if a node has an initial level of priming of 10, multiplication of priming will reach threshold (100) in less than 4 ms.

SEQUENTIAL RULES

Connections between sequence nodes represent serial-order rules that determine order of activation when two or more sequence nodes have received simultaneous priming. Serial-order rules represented by sequence nodes in the sentential system are termed syntactic rules, and serial-order rules represented by sequence nodes in the phonological system are termed phonological rules (see Table 3.1 for typical examples). The sequence nodes ADJECTIVE and NOUN, for example, are connected in such a way as to represent the syntactic rule that adjectives precede nouns in English noun phrases. Similarly, the sequence nodes INITIAL CONSONANT GROUP and VOWEL GROUP are connected in such a way as to represent the phonological rule that initial consonants in a syllable precede the vowel and final consonants.

An inhibitory connection is a simple means of achieving this order relation among sequence nodes. Under this proposal, the first to be activated of a pair of sequentially organized sequence nodes inhibits the next to be activated whenever both are simultaneously primed. This inhibitory connection therefore enables the first to be activated to become activated first under the most-primed-wins principle. Following activation of the first sequence node, the next to be activated is released from inhibition and can then become activated under the most-primed-wins principle. For example, ADJECTIVE inhibits NOUN and dominates

TABLE 3.1. Examples of serial-order rules in the phonological and sentential systems.

Example Serial-Order Rules	Example Instances
Phonological System	
Initial consonant group + vowel group	*str + and*
Initial fricative + initial stop + initial liquid	*s + t + r*
Vowel + glide + liquid	*o + w + l*
Final nasal + final stop	*n + d*
Sentential System	
Noun phrase + verb phrase	*the theory + enabled extensive progress*
Verb + noun phrase	*enabled + extensive progress*
Determiner + noun	*the + theory*
Adjective + noun	*extensive + progress*

in degree of priming whenever ADJECTIVE and NOUN receive simultaneous priming. However, once ADJECTIVE has been activated and its priming returns to resting level, NOUN is released from inhibition and dominates in degree of priming, thereby determining the sequence (adjective + noun) for this and any other noun phrase containing an adjective and a noun.

Sequential rules such as adjective + noun bear a surface resemblance to phrase structure rules of Chomsky (1957), such as noun phrase → noun + adjective, where the arrow stands for "is rewritten as." Both types of rules are nontransformational, for example, and refer in this example to identical sentential domains or syntactic categories, the set of all adjectives and nouns. There are many differences, however (D. G. MacKay, 1974). For example, there is no sense in which the sequence node NOUN PHRASE is "rewritten" as NOUN + ADJECTIVE in the node structure theory. Rather, the lexical content nodes connected to this particular noun phrase node simultaneously prime their respective sequence nodes, which happen to be NOUN and ADJECTIVE.

The node structure theory also postulates new processes and sequential domains that were unforeseen in phrase structure grammars, such as sequential domains for discourse nodes (D. G. MacKay, 1985), for morphological nodes (D. G. MacKay, 1973b), and for phonological nodes, as in *initial consonant group + vowel group*. Moreover, the seeming equivalence of some phrase structure categories and sequential domains in the examples discussed so far is fortuitous. Sequential domains such as *adjective* and *noun* are in fact much more fine grained than I have so far discussed. For example, the domain of *green* and *red* must be *color adjective*, and the domain of *frequent* and *fast* must be *temporal adjective*, rather than just *adjective*. Although all adjectives precede nouns in English, more restricted domains, such as *color adjective* and *temporal adjective*, are necessary in order to ensure the appropriate sequencing among different types of adjectives. Thus, a sequential rule such as *temporal adjective + color adjective* is required in order to produce the usual sequence when temporal and color adjectives are conjoined. For example, we normally say "frequent red lights," instead of "red frequent lights." Without domains such as *color adjective* and *temporal adjective*, we would have no mechanism for producing the preferred rather than nonpreferred adjective order. Factors such as emphasis can of course alter this preferred or neutral order.

Subdivisions within the domain that I have represented here as nouns are likewise necessary, and for similar reasons. The only constraint on sequential domains in the theory is that their corresponding sequence nodes must be "called up" by means of direct connections from content nodes. Is it possible to obey this constraint in describing the overall system of sequential rules for English or any other language? What is the full set of sequential domains for producing English? These questions currently lack conclusive answers (but see Gazdar, 1981).

However, the way that content nodes call up sequence nodes in parallel and the way that sequence nodes, after interacting among themselves, activate content nodes seems to be exactly the sort of mutually interacting process that is needed for explaining recent demonstrations of an influence of syntax on word recognition (Isenberg, Walker, Ryder, & Schweikert, 1980) and on word selection (Bock & Warren, 1985) and vice versa (Bock & Warren, 1985).

THE ACTIVATION OF SEQUENCE NODES

Timing nodes are the mechanism for activating sequence nodes and are connected with sequence nodes in the same one-to-many way that sequence nodes are connected with content nodes. Separate mechanisms therefore determine the form, sequence, and timing of behavior in the node structure theory, but timing is more closely related to sequencing than to the form of behavior; timing nodes are directly connected to sequence nodes, but not to content nodes.

Timing nodes become activated according to an endogenous rhythm, and timing nodes at different levels have different endogenous rhythms (Chapter 5). Following each activation, timing nodes multiply the priming of the sequence nodes connected to them, activating the most primed one on the basis of the most-primed-wins principle. By determining how rapidly the sequence nodes become activated, timing nodes therefore determine the rate of the output, a topic taken up in detail in Chapter 5, along with other aspects of the temporal organization of perception and action.

Functional Relationships Between Sequence and Content Nodes

Sequence nodes perform three main functions in relation to content nodes: They organize the content nodes into domains, they activate the most primed content node in a domain, and they determine the serial order in which the content nodes become activated.

THE ORGANIZING FUNCTION

A domain consists of a set of nodes that all become activated by means of the same mechanism. For example, the hypothetical sequence node TEMPORAL ADJECTIVE is connected to, and serves to activate, the many content nodes in the domain *temporal adjective*, thereby organizing these nodes together into a single domain. However, domains reflect a functional relationship among nodes. Domains should not *in general* be thought of as sets of nodes aggregated into nonoverlapping anatomical areas of the brain. Although such localization may actually occur in lower level systems (D. G. MacKay, 1985), one and the same content node in a higher level system can receive connections from several sequence nodes and thereby occupy several domains. For example, consider the sentential system representation of the many English words that can be used with identical meaning as either nouns or verbs, for example, *practice*. A single content node, *practice*(noun, verb), represents this word and becomes activated by either NOUN or VERB. Because dual-function content nodes, such as this one, simultaneously occupy more than one domain, higher level domains must overlap, at least to some extent, at the neuroanatomical level.

From a functional point of view, a sequential domain consists of a set of nodes, all of which have the same sequential function. Thus, the domain of content nodes representing nouns all have the same sequential properties or privileges of occurrence in English sentences. Similarly, the domain of nodes representing vowels all have the same sequential properties, or privileges of occurrence, in the syllables of English or any other language.

The Triggering Function

The second function of sequence nodes is to activate whichever content node has greatest priming in its domain. This most-primed-wins principle for activation follows directly from the nature of connections between sequence and content nodes. Once a sequence node becomes activated, it automatically and simultaneously multiplies the priming of the entire domain of content nodes connected with it, so that their level of priming increases rapidly over time. However, the intended-to-be-activated node in the domain has just received priming "from above," because its superordinate node (see Figure 3.1) has just been activated. Being most primed (usually), this primed-from-above node reaches threshold sooner than other "extraneous" nodes in its domain and becomes activated.

Once a content node becomes activated, its sequence node must return quickly to resting level, because content nodes have a return connection to their sequence node, which could cause reverberatory reactivation. Thus, an activated content node must quench, or inhibit, rather than further prime, its corresponding sequence node, so that only one content node becomes activated at any one time. Without being quenched, a sequence node could simultaneously activate several nodes in its domain, causing a potential breakdown in behavior. Quenching requires a threshold mechanism, which if exceeded, causes content nodes to inhibit rather than prime their sequence nodes.

The Sequencing Function

It is important to understand that priming is fundamentally nonsequential. An activated content node primes all of its connected nodes at the same time. Sequence nodes are needed to impose the sequence of activation, and thereby determine the appropriate temporal sequence for words, segments, and muscle movements in the final output. For example, activating *practice*(noun, verb) simultaneously primes *prac*(stressed syllable) and *tice*(unstressed syllable), and as discussed above, the connections between sequence nodes represent which one comes first, enabling the correct sequence to be generated in the final output.

A SIMPLIFIED EXAMPLE

To illustrate how timing and sequence nodes interact to determine whether, when, and in what order content nodes become activated in everyday speech production, consider how the words *frequent* and *practice* might become sequenced in the noun phase "frequent practice." Figure 3.2 illustrates the

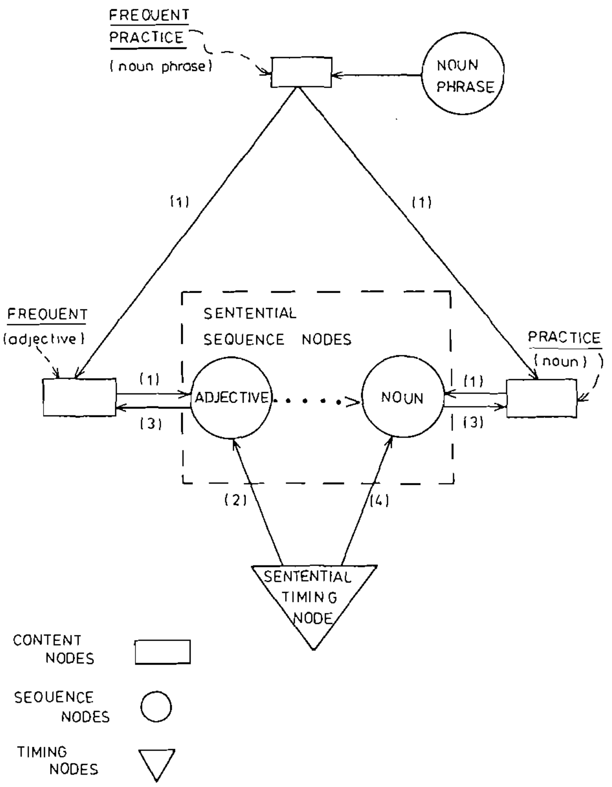

FIGURE 3.2. The order of top-down processes (in parentheses) underlying activation of content nodes (in rectangles), sequence nodes (in circles), and the sentential timing node (in triangle) for producing the noun phrase *frequent practice*. (From "The Problems of Flexibility, Fluency, and Speed-Accuracy Trade-Off in Skilled Behavior" by D. G. MacKay, 1982, *Psychological Review*, *89*, p. 492. Copyright 1982 by American Psychological Association, Inc. Reprinted by permission.)

hypothetical content nodes (in rectangles), sequence nodes (in circles), and timing node (triangle). Unbroken lines in Figure 3.2 are excitatory, the broken lines represent the quenching mechanism, and the dotted lines represent the inhibitory relationship between sequence nodes. Similar connections and processes are postulated for all sequentially organized mental nodes.

Assume that the node representing the sentential concept, *frequent practice*(noun phrase) has received top-down priming and is activated first in the same way as any other nodes, such as the ones described below. Activation of *frequent practice*(noun phrase) causes simultaneous priming of *frequent*(adjective) and *practice*(noun), which immediately pass on second-order priming to their hypothetical sequence nodes, ADJECTIVE and NOUN. The inhibitory link between ADJECTIVE and NOUN temporarily reduces the priming level for NOUN, so that ADJECTIVE is most primed and becomes activated under the most-primed-wins principle following the first pulse from the timing node. ADJECTIVE therefore multiplies the priming of every content node in its *adjective* domain, and the one with the most priming in the domain reaches threshold soonest and becomes activated (the most-primed-wins principle). The node with the most priming in the *adjective* domain will of course usually be *frequent*(adjective), which has just recently been primed by *frequent practice*(noun phrase).

Once activated, *frequent*(adjective) "quenches" its sequence node; that is, it quickly reduces the activity of ADJECTIVE to resting level, thereby ensuring that one, and only one, content node in the domain becomes activated at any given time. Quenching of ADJECTIVE releases the inhibition on NOUN, which therefore dominates in degree of priming within the domain of sentential sequence nodes, and NOUN becomes activated under the most-primed-wins principle following the next pulse from the sentence timing node. NOUN therefore multiplies the priming of the entire domain of *noun* nodes, but having just been primed, *practice*(noun) has more priming than any other node in the domain and becomes activated under the most-primed-wins principle.

This example gives a flavor of the complexity of the processes that must underly sequencing in a skilled behavior such as speech production. As already noted, however, I have chosen a simple example. It focuses on only two nodes among the millions of nodes that are relevant to the way words combine to form the grammatical sentences of English. And I have simplified this simple example even further by representing the domain of, say, the word *frequent*, as *adjective*. As discussed above, the domain of *frequent* must be *temporal adjective*, rather than just *adjective*.

Differences Between Sequence and Content Nodes

How do sequence and content nodes differ? Although sequence and content nodes are similar in some respects, they differ in others, and the differences summarized in the following section are important for understanding the theory.

FUNCTIONAL DIFFERENCES

Two basic functions that differentiate sequence nodes from content nodes are activation and sequencing. Sequence nodes can cause their connected nodes to become activated but content nodes cannot. No matter how long a content node primes a connected node, activation cannot occur without the help of the triggering mechanism or sequence node specific to the primed node. With respect to the sequencing function, content nodes cause simultaneous, that is, cotemporal or nonsequential activity in their connected nodes. Priming from an activated content node is invariably transmitted to all of its connected nodes at the same time. Only sequence nodes give rise to sequential activity by activating content nodes in a predetermined order.

QUANTITATIVE DIFFERENCES

Both content and sequence nodes are organized into domains, receive priming from connected nodes, and are activated by multiplication of priming under the most-primed-wins principle. However, there are *quantitative* differences in the connections and domains for sequence versus content nodes. A content node typically connects with a single sequence node and with only two or three content nodes, which usually occupy different domains. In contrast, a sequence node connects with up to a thousand content nodes, which always occupy the same domain. For example, NOUN only connects with content nodes in the domain *noun*. A sequence node also connects with other sequence nodes, but only ones in its own domain. For example, NOUN only connects with other sequence nodes in its own (sentential sequence) domain.

Content and sequence nodes also differ greatly in number. Because different sequence nodes by and large connect with different content nodes and because each sequence node connects with many (up to a thousand) content nodes, content nodes must outnumber sequence nodes by a ratio of up to a thousand to one.

The number of domains that content versus sequence nodes are organized into also differs by at least an order of magnitude. Content nodes are organized into as many domains as there are sequence nodes. In the case of speech production, there are hundreds of sequence nodes and, therefore, hundreds of content node domains. In contrast, sequence nodes for speech seem to be organized into only about three domains: sentential sequence nodes, phonological sequence nodes, and (speech) muscle movement sequence nodes.

In summary, sequence nodes can be considered semispecific rather than nonspecific activating mechanisms. Because sequence nodes connect with, and activate, a specific and limited set of content nodes in domains such as *color adjective*, they exhibit some degree of specificity. However, sequence nodes also exhibit considerable nonspecificity, and certainly more nonspecificity than content nodes, because they *always* connect with, and activate, a category of content nodes rather than just one.

DIFFERENCES IN CONNECTIONS

Connections between one content node and another differ from connections between one sequence node and another. Connections from one content node to another are simple (nonmultiplicative) and excitatory in nature and usually involve nodes in different domains. For example, a noun phrase node usually connects with nodes in domains such as *adjective* and *noun*. In contrast, connections between sequence nodes are inhibitory in nature and always involve nodes within the same domain. For example, the sequence node ADJECTIVE inhibits NOUN within its own domain of sentential sequence nodes for producing English.

Connections from content to sequence nodes also contrast with connections from sequence nodes to content nodes. Sequence nodes only send a multiplicative connection to their connected content nodes, while content nodes send a simple excitatory connection and a quenching connection to their connected sequence node(s).

CENTRALITY TO A SYSTEM

Sequence nodes can be considered more central to a system than timing and content nodes. For example, sequence nodes receive connections from all three types of nodes: timing nodes, content nodes, and other sequence nodes. Content nodes, although essential for all instances of behavior, are structurally more peripheral than sequence nodes and only receive connections from sequence nodes and other content nodes. Timing nodes are likewise essential, at least for skilled behavior, but are the least central type of node in a system. A timing node receives connections from neither sequence nor content nodes within its own system.

Implications of the Theory

The node structure theory was designed to incorporate the general requirements for theories of sequencing outlined in the introductory section of this chapter, and it should come as no surprise that the theory explains the phenomena discussed there. I therefore focus on only two of these phenomena, in order to illustrate some implications of specific applications of the theory.

Production Onset Time

Recall that production onset time (the time to begin to produce a preplanned behavior) is shorter when the behavior consists of a single component than when it consists of a sequence of components. The node structure theory explains this finding as due to the time required to prime and activate the nodes preceding the first muscle movement node in the sequence. Some of these activated nodes are sequence nodes, and onset time also depends on the interactions between sequence nodes, that is, the number of serial-order rules that must become activated (or, equivalently, sequential decisions that must be made; see Figure 2.5) prior to activating the first muscle movement nodes.

Temporal duration at the surface level is irrelevant to production onset time under the theory, so that, despite the large differences in surface duration, only small increases in onset time can be expected for a one- versus two-syllable word, for a one- versus two-sentence paragraph and for a one- versus two-topic preplanned lecture (all other factors except output duration being equal). Even the number of preplanned components per se is irrelevant to onset time. As already noted, production onset time for, say, *pain* and *paint* should be equivalent under the theory (all other factors except length being equal), because the same number of nodes and serial-order rules must become activated before activating the first muscle movement node for /p/. The *nt* node and the serial-order rule for n + t in *paint* only become activated after the first muscle movement node has been primed and activated. Here the extra nodes and sequential operations leave onset time unchanged and add only to the overall time to produce the word. This prediction contrasts sharply with predictions from other theories (e.g., Klapp, 1979), which view the temporal duration of a preplanned sequence as the critical determinant of production onset time.

The node structure theory also provides a coherent account of Klapp and Wyatt's (1976) onset time data for sequences of finger movement. Recall that production onset times are longer for Morse code sequences beginning with a *dah* than a *dit*. A *dit* requires three hierarchically organized content nodes above the muscle movement level (see Figure 3.3): the highest level "*dit node*," and two subordinate nodes, one representing the press, and the other the lift for releasing the key. In support of this representation, Wing (1978) showed that the mechanisms for pressing and lifting the finger from a key are independent. A careful analysis of the timing characteristics of the lift and press components in an experiment involving repetitive finger tapping indicated that lifting the finger from the key is not triggered by the preceding press or vice versa. As Wing (1978) points out, the independence of these components is also required to explain the blocks or temporary pauses that sometimes occur in rapid finger tapping and other repetitive activities (see also Glencross, 1974). These blockages suggest that the independent press and lift components fall progressively out of phase until they become activated simultaneously rather than sequentially at the muscle movement level, so that no movement can occur.

Note, however, that a *dah* is more complex than a *dit* by any analysis: a *dah* requires an additional node for representing the fact that the key must be held in contact with the terminal, plus a timing mechanism for specifying the duration of this contact phase. (Figure 3.3 leaves out the timing mechanism but represents one possible relation between the content nodes for a *dah*, although not the only possible relation given our current state of knowledge about Morse code.) The longer onset time for *dah* (in either first or second position) may therefore

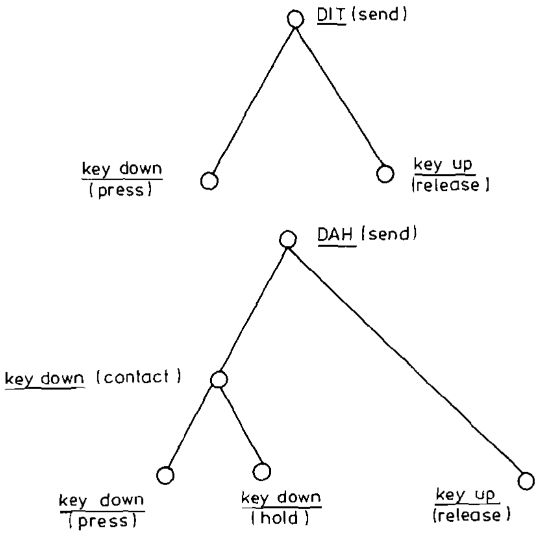

FIGURE 3.3. The structure of content nodes for a beginner producing *dit* versus *dah* in Morse code (D. G. MacKay, 1985).

reflect the greater number of content nodes, serial-order rules, and timing pulses that must be activated to begin producing a *dah*.

The theoretically more challenging finding concerns the longer initiation times for sequences containing different components (i.e., *dah-dit* and *dit-dah*) than for sequences containing identical components (i.e., *dah-dah* and *dit-dit*). Under the node structure theory, this finding reflects a difference in the mechanisms for repeating an element versus sequencing different elements. A repeated component requires a simple repeat mechanism, without any sequential decision (interaction between sequential nodes representing a sequential rule), whereas sequencing different components requires a (time-consuming) sequential decision involving interactions between two or more sequence nodes. Moreover, in this particular case, the two sequential rules conflict with one another (*dit* + *dah* for one sequence and *dah* + *dit* for the other) and cannot be called up in advance, adding further to the time required to resolve which element comes first under the most-primed-wins principle.

The Sequential Class Regularity

The sequential class regularity occurs in speech errors when one linguistic component inadvertently substitutes another. The substituted and substituting components almost invariably belong to the same sequential class. For example, over 99% of the word substitution errors in Stemberger's (1985) corpus obeyed the sequential class regularity. Nouns substituted with other nouns, verbs with verbs, and not with, say, nouns or adjectives. As noted earlier, this regularity also holds for other levels of speech production: for substituted morphological components (prefixes substitute with other prefixes, suffixes with other suffixes, and never prefixes with suffixes); for substituted syllabic components (initial consonant clusters substitute with other initial clusters, final with final, but never initial with final); and for substituted segmental components (vowels substitute with vowels, consonants with consonants, and never vowels with consonants).

The sequential class regularity also holds for other types of errors: anticipations, perservations, and transpositions (although somewhat less strongly; see the exceptions discussed in the following section). As Meringer and Mayer (1895) pointed out, a common mechanism almost certainly underlies all three classes of errors, and the most-primed-wins principle is exactly the sort of common underlying mechanism that is needed. Errors occur under the node structure theory whenever an "intended" node has less priming than some other "extraneous" node in its domain when the activating mechanism is applied. Because an activating mechanism also applies to a particular domain or sequential class, this means that substituted components at every level in the system will belong to the same sequential category. For example, if *crotch*(noun) acquires greater priming than *watch*(noun), for whatever (e.g., Freudian) reason, *crotch* will substitute for *watch* when the activating mechanism is applied to the noun domain. However, the noun *crotch* will never substitute with a verb, even a phonologically similar verb such as *botch*, because the corresponding activating mechanisms cannot be applied simultaneously to both the noun and the verb domains. Because of the most-primed-wins principle, only one sequence node can be applied at a time; either NOUN or VERB can become activated, but not both simultaneously. Of course, this is not to say that one initial consonant, /b/, cannot substitute for another, /w/, to cause the substitution of *botch* for *watch*. When this occurs, however, the error involves not a word substitution, but a phonological substitution which obeys a sequential class regularity of its own: initial consonants substitute with initial consonants.

The node structure theory predicts, further, that the sequential class phenomenon will hold statistically for *all* types of errors, including also nonsequential errors such as blends and malapropisms. The reason is of course that an activating mechanism can only activate and misactivate nodes within the same sequential domain as the appropriate or intended-to-be-activated node. A similar regularity should also hold for errors in other highly skilled behaviors, even those involving very different sequential classes, such as typing, where the two hands seem to make up one domain; the different types of strokes (horizontal, vertical, lateral) make up a second; and the homologous fingers of the two hands make up a third class of domains (Grudin, 1981).

EXCEPTIONS TO THE SEQUENTIAL CLASS RULE

About 20% to 30% of Stemberger's 1985 corpus of sequential errors (anticipations, perseverations, and transpositions) failed to preserve sequential class, but these relatively infrequent exceptions to the sequential class rule often display regularities of their own that must be explained in theories of sequencing. Consider the following examples from Fromkin (1973), "She was waiting her husband for" (instead of, "waiting for her husband") and "I don't want to part this book with" (instead of, "to part with this book"). Errors such as these pose three questions: Why do they violate the sequential class rule? (Both errors involve a noun phrase changing places with a verb particle.) Why are they so rare? And why do they result in a sequence (verb + noun phrase + verb particle) that is appropriate for other expressions such as "She called the man up"? The node structure answer to these questions is that these errors reflect misapplication of a serial-order rule. Connections with the wrong sequence nodes (and therefore the wrong serial-order rules) have been formed, which results in the wrong order in the output. For example, a connection to *for*(particle) rather than, or even in addition to, a connection to *for*(preposition) could result in an error such as "waiting her husband for."

In short, the theory distinguishes two general classes of speech errors: *priming* errors and *rule* errors (see also Stemberger, 1985). Priming errors obey the sequential class phenomenon, and occur whenever an intended node has less priming than some other node in its domain at the time when the most-primed-wins activating mechanism is applied. Rule errors violate the sequential class phenomenon and occur whenever an inappropriate serial-order rule is called up or primed. Rule errors are relatively rare, because forming incorrect connections between nodes is relatively rare.

Of course the node structure theory doesn't yet provide a full or conclusive account of all aspects of speech errors. One of the many outstanding questions is why higher level nodes in a system are less prone to error than lower level nodes in the same system (Dell, 1985a). In the sentential system, for example, phrase nodes participate in fewer errors than do lexical nodes, and in the phonological system, syllable nodes participate in fewer errors than do segment or feature nodes. The answer is certainly *not* that higher level content nodes for phrases and syllables do not exist. After all, these units *sometimes* participate in errors. The answer suggested by the node structure theory is related to speed–accuracy tradeoff. Nodes at higher levels in a system are less prone to error because they are activated at slower rates than lower level nodes in a system. For example, syllables may be less prone to error than segments or features because segments and features are produced much faster than syllables in the phonological system. A new syllable node is activated every, say, 500 ms, whereas a new segment node or set of feature nodes is activated every, say, 150 ms, allowing less time for priming to summate, and thereby increasing the probability of activating the wrong segment node under the most-primed-wins principle.