Chapter 13

Stuttering as a Sequencing and Timing Disorder

Donald G. MacKay and
Maryellen C. MacDonald

The present chapter takes up the problem of stuttering where Van Riper (1982) left off. Van Riper (p. 45) defined stuttering as a disruption of the simultaneous and successive programming of muscular movements required to produce a speech sound or its link to the next sound in a word. Anticipation of this programming difficulty can then cause struggle and avoidance reactions which are secondary, variable, and learned. However, the primary difficulty lies in the programming of sequence and timing, and Van Riper summarized several sources of evidence for this basic thesis. For example, asynchronies or lags have been observed in all of the speech muscles of stutterers, not only during overt instances of stuttering but also during their seemingly fluent speech (Zimmerman, 1980). The temporal coordination of voice, respiration, and articulation is apparently disrupted during fluent as well as nonfluent utterances of stutterers. Stutterers are also less able to repeat the temporal pattern of a sentence or sequence of finger taps than nonstutterers (Cooper & Allen, 1977), as if their neural clocks are less accurate or more susceptible to mistiming.

Van Riper next showed how timing disruptions might account for many of the basic phenomena observed in research on stuttering. An example is the rhythm effect, where fluency is enhanced when a stutterer speaks in time with a metronome or any other rhythmic stimulus (visual, auditory, or tactile), provided the rhythm is not abnormally fast. According to Van Riper (1982) externally generated rhythm may help facilitate the timing of motor patterns which are prone to asynchrony in stutterers. Fluency is likewise improved when stutterers sing because musical rhythm may help facilitate timing of the syllables corresponding to the notes.

Despite this evidence favoring his hypothesis (see also Perkins, Bell, Johnson, & Stocks, 1979), Van Riper concluded his book on a pessimistic note and gave three reasons for his pessimism. First, he found it "very difficult to evaluate the degree of support all of this scattered and often indirectly focused research lends to the position that stuttering involves a disruption in timing." Second, lacking a theory of how timing is achieved in fluent or unstuttered speech, Van Riper was unable to specify the nature and cause of the hypothesized timing difficulty. Finally, Van Riper was concerned that viewing stuttering as a timing disorder may be incompatible with data indicating that stuttering is related to the processing of auditory feedback.

Expressions of discontent similar to Van Riper's are rampant in recent literature. Preus (1981) calls the state of stuttering research "deplorable" (p. 13), and Bloodstein (1969) finds stuttering theories either descriptive in nature or so vague as to be unhelpful. As Bloodstein points out, to call stuttering a perseverative response, a symbolic sucking activity, or a miniature convulsion, only describes rather than explains it. Similarly, attributing stuttering to anxiety, stress, or delayed myelinization of cortical neurons is theoretically unhelpful unless a detailed causal explanation can be provided for at least one specific, real-time example of the moment of stuttering.

THE METATHEORY UNDERLYING STUTTERING RESEARCH

The present chapter addresses all of these concerns and another more general one which is relevant to virtually all research and theories of stuttering to date. It concerns the metatheory underlying past stuttering research. Under this metatheory, stuttering can be studied by itself, independent of both data and theories on how normal, error-free speech is achieved. In short, this metatheory views stuttering research as a field unto itself with its own special methodology, phenomena, and theories.

This metatheory explains why few studies of stuttering have made attempts to integrate the findings from normal speakers with those from stutterers (Garber & Siegel, 1982) and why studies of stuttering have proceeded in virtual isolation from the remainder of psychology and speech science. Moreover, this metatheory has provided a serious obstacle to our understanding of stuttering and its relationship to other speech errors. Any theory developed under this metatheory is, at best, a stab in the dark. Constructing a separate theory of stuttering is analogous to constructing a separate theory of backfires for explaining why cars sometimes emit explosive noises from their exhaust systems. To really explain backfires one must begin with an understanding of the principles of internal combustion which govern the normal functioning of an automobile engine. Similarly, to explain stuttering, one must begin with

an understanding of the mechanisms underlying the production of error-free speech (see MacKay, 1969b, 1970a).

Of course there are good reasons why stuttering research has adopted a "stab-in-the-dark" metatheory and has proceeded independently of theories of normal speech production. The field has an understandable desire to provide immediate relief for stutterers, and theories of normal speech production from which to derive an explanation of stuttering, or any other class of speech errors, have been slow in coming and "woefully sketchy" (Van Riper, 1982). The reason for such sketchiness lies in the number and complexity of the processes that must be timed and sequenced in the control of fluent speech. First, there are many different levels or systems of control: the sentential system for controlling the sequencing of words in sentences; the phonologic system for controlling the sequencing of syllables and phonemes within words; and the muscle movement system for controlling and coordinating the laryngeal, respiratory, and articulatory muscles for producing speech sounds.

Moreover the processes within each of these systems are extremely complex. Considering only the lowest level speech musculature, over 100 different muscles may be involved in producing a single word, and each must get its appropriate nervous impulses at the required moment in the sequence if the word is to be spoken without disruption. At the time Van Riper (1982) wrote, both the normal events and the disruption of these events during the moment of stuttering were considered so complex as to preclude detailed theoretical description. Woeful sketchiness was inevitable.

However, recent years have seen significant advances in the understanding of processes underlying the sequencing and timing of speech, especially within the sentential and phonologic systems (see MacKay, 1982), and there is reason to believe that similar mechanisms may play a role within the muscle movement system as well. We will begin by outlining a general theory of speech production and then examine how and why the postulated mechanisms for timing and sequencing fluent speech may become disrupted during moments of stuttering. We shall focus especially on the question of where in the speech production system stuttering originates and on how this theory may account for phenomena such as the effects of adaptation and altered feedback on the occurrence of stuttering.

A GENERAL THEORY OF SPEECH PRODUCTION

The basic components underlying motor control in the theory are content nodes, each consisting of one or more neurons (MacKay, 1982).

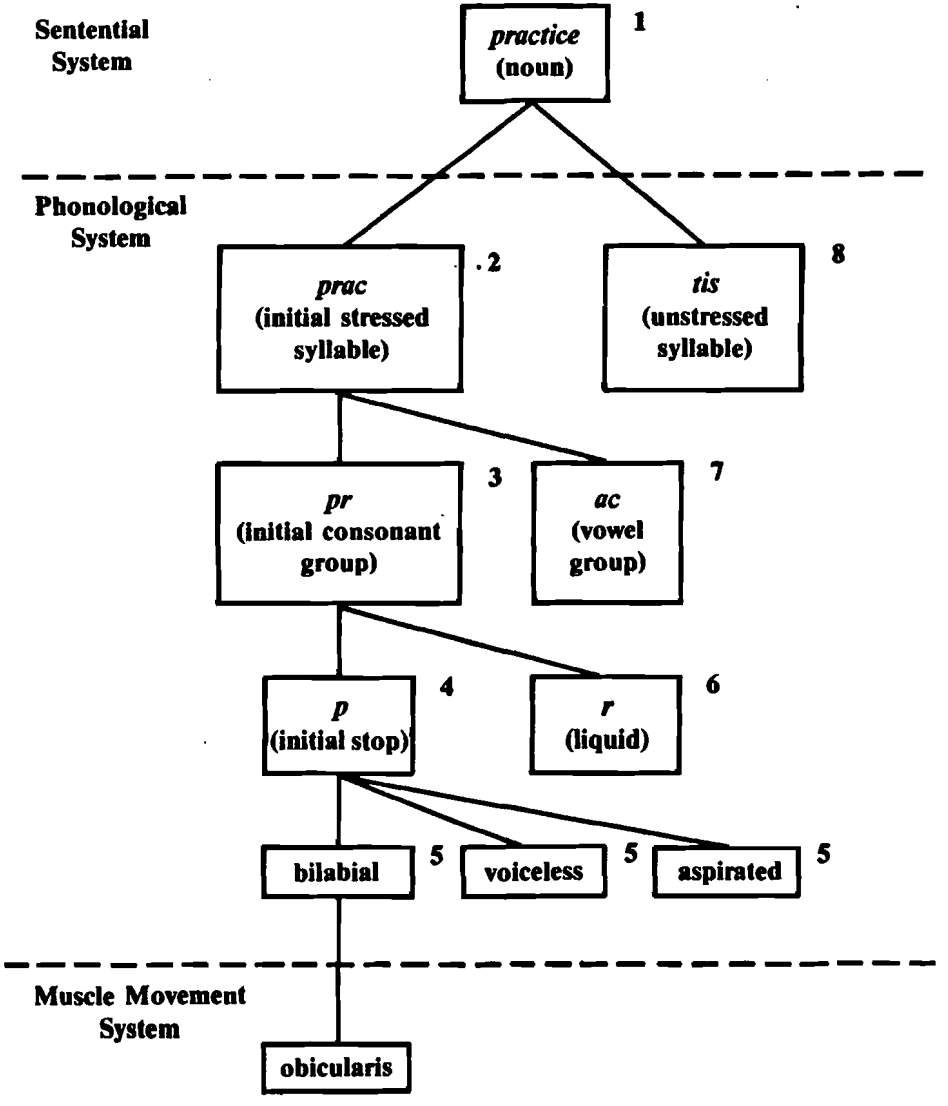
In the case of speech production, the content nodes are organized into three independently controllable systems: the muscle movement system, the phonological system, and the sentential system. Content nodes within the muscle movement system represent muscle-specific patterns of movement involving the respiratory system, larynx, and articulatory organs such as the tongue, velum, and lips. Content nodes within the phonologic and sentential systems represent not specific muscles but cognitive units for controlling the movements making up a preprogrammed sequence such as a word or a phrase.

The motor program for words

To keep matters simple, we focus on the relatively small number of nodes making up the motor program for producing a single word. The components of a word can be represented by a hierarchy of interconnected content nodes (see MacKay, 1972, and Treiman, Salasoo, Slowiaczek, & Pisoni, 1982, for detailed evidence bearing on the organization of nodes underlying words). The highest level content node represents the concept underlying the word. By way of illustration, consider the noun, *practice*, which is represented by a single content node in the sentential sytem (see Figure 13-1). When *practice* (noun) becomes activated, its connected syllable nodes *prac* (initial stressed syllable) and *tis* (unstressed syllable) become primed or readied for activation. Unlike activation, priming varies in degree and summates over time up to some asymptotic level. Also unlike activation, priming is automatic and parallel in nature and requires no special triggering mechanism to determine when or in what sequence it occurs. However, activation must occur in a predetermined sequence if the word is to be produced without error. In the example under consideration, each node will be activated in the order shown in Figure 13-1.

Syllable nodes such as *prac* (initial stressed syllable) are part of the phonologic system and are connected to nodes representing phonologic compounds such as *pr* (initial consonant group) and *ac* (vowel group). These in turn are connected to nodes representing phonemes such as *p* (initial stop) and *r* (liquid). A phoneme node such as *p* (initial stop) is connected to distinctive feature nodes which are connected in turn to a hierarchy of muscle movement nodes representing patterns of movement for the various muscles such as the obicularis oris muscle for the lips (see Figure 13-1). The structure of connections between nodes in the muscle movement system (unlike the phonologic and sentential systems) is currently unknown.

FIGURE 13-1. The organization of selected content nodes within the sentential, phonological, and muscle movement systems for producing the noun *practice*. Numbers indicate order of activation of the units shown and the domain or sequential class of each node appears in brackets. See text for explanation.



Nodes are hypothesized to share three dynamic properties which are relevant to the occurrence of stuttering and other speech errors: priming, activation, and linkage strength.

Priming. Priming is an excitatory input that active nodes pass on to nodes connected to them. The priming a node receives summates over time up to an asymptotic level, at which point a node is fully readied for activation.

Activation. Activation is the highest level of activity of a node. It is all or none in nature and is self-sustained for a specifiable period of time. Activation of the lowest level muscle movement nodes is necessary for behavior to occur, and the order and timing of activation of these nodes determines the sequence and timing of activity in the final output.

Linkage strength. Linkage strength is a long-term characteristic of the connection between nodes which determines the asymptotic level and rate at which a connection passes priming from one node to another. The main variable influencing linkage strength is practice: the frequency with which a particular connection has been activated in the past. Increased linkage strength yields lower probabilities of error for a given rate of speech (see MacKay, 1982).

The sequential activating mechanism: Sequence nodes

Activation of nodes that have been primed is achieved by a nonspecific activating mechanism which is responsible for activating an entire set or domain of nodes. Sequence nodes (capitalized in the examples to follow) are posited as a nonspecific activating mechanism for activating content nodes in proper serial order. For example, the sequence node NOUN is responsible for activating *practice* (noun) and every other node representing noun concepts. At any one time, however, a sequence node activates only one content node—the one with the greatest degree of priming. This "most-primed-wins principle" applies to the activation of every node in every system, including the sequence nodes themselves. In the case of content nodes, the node with the greatest degree of priming will normally be the one that has just been primed via its connection to an activated node which is superordinate in the hierarchy (see Figure 13-1). Sequence nodes are an independently stored set of nodes within each system, and the connections between them represent the sequential constraints for the classes of content nodes in question. For example, the sequence nodes INITIAL STOP and INITIAL LIQUID for producing the initial consonant cluster (*pr*) in *practice* are connected in such a way as to represent the fact that stops invariably precede liquids in initial clusters in English. This precedence relation among sequence nodes is achieved by an inhibitory connection. Thus, INITIAL STOP inhibits INITIAL LIQUID and dominates in degree of priming whenever these two sequence nodes are simultaneously primed by content nodes. As a consequence, INITIAL STOP can be activated first under the most-primed-wins principle as the most primed node in the domain of phonological sequence nodes.

The temporal activating mechanism: Timing nodes

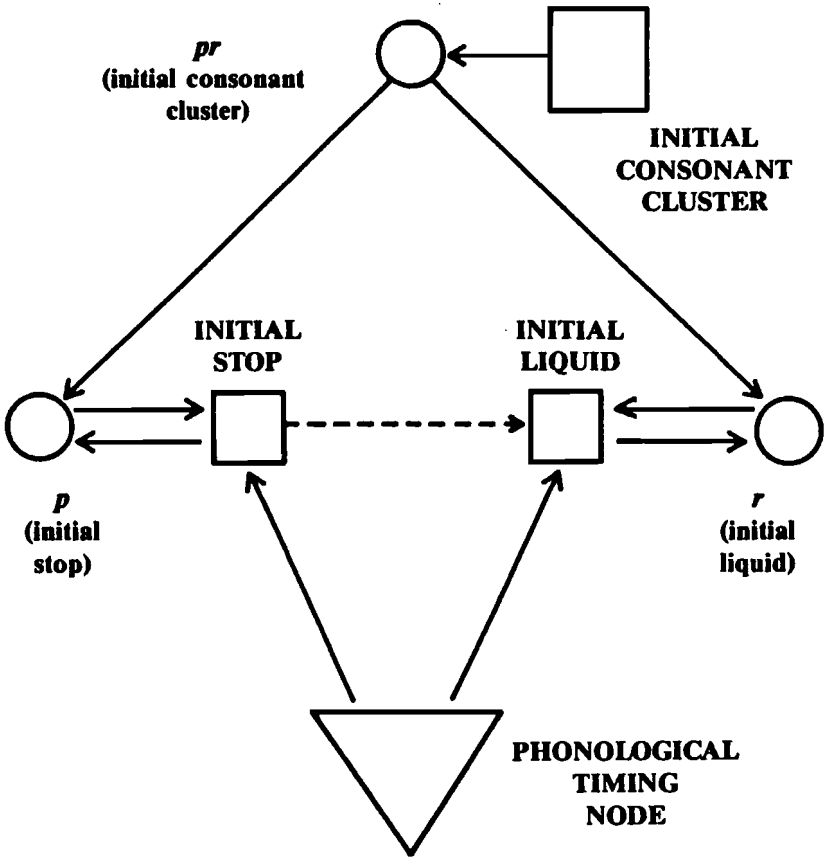
Timing nodes represent the components of an internal clock which determines when the sequence nodes become activated. Timing nodes bear the same relation to sequence nodes that sequence nodes bear to content nodes. A sentence time node is connected to sequence nodes for activating the content nodes coding the components of sentences. A phonologic time node is connected to sequence nodes for activating the content nodes coding the components of syllables. And a muscle time node is connected to sequence nodes for activating the content nodes controlling muscle movements within the laryngeal, respiratory, and articulatory systems. Timing nodes, therefore, constitute the underlying basis for organization of the nodes into the three systems discussed in the introduction. Each timing node sends out pulses at specifiable intervals, but the mean pulse rate for the three timing nodes differs. For example, the phonologic time node must generate more pulses per second than the sentence time node since phonemes are produced faster than words (by a factor of about 5 on the average). However, the three timing nodes are coupled and operate in a correlated manner when simultaneously active: If the sentence time node is speeded up, the phonologic and muscle time nodes are speeded up proportionally. Whenever a timing node becomes activated it simultaneously primes the entire set of sequence nodes connected to it, and this priming summates quickly over time so that the sequence node with the greatest degree of priming reaches threshold and becomes activated.

Finally, the timing nodes for the three systems can be independently controlled. If only the sentence time node becomes activated, propositional thought without internal (phonological) speech will occur. If both the sentence and phonological time nodes become activated, internal speech will occur without overt movement of the speech musculature (even though the appropriate muscle movement nodes will have been primed). Only when all three timing nodes are simultaneously active will speech take place.

An example

To illustrate how the timing and sequence nodes interact to determine whether, when, and in what order content nodes become activated, we

FIGURE 13-2. Processes underlying activation of three content nodes in the phonological system (circles), their corresponding sequence nodes (rectangles), and timing node (triangle).



propose to focus on the first two phonemes in *practice*. Identical processes are assumed to underlie activation of nodes at every level in any system (see MacKay, 1982, for a more detailed account of these processes).

The nodes and connections between them which are relevant to this example are shown in Figure 13-2. Unbroken connections are excitatory, and the dotted connection between sequence nodes (rectangles) is inhibitory. The node representing the superordinate component *pr* (initial consonant cluster) is activated first. This simultaneously primes two subordinate content nodes: *p* (initial stop) and *r* (initial liquid) which in turn prime their corresponding sequence nodes INITIAL STOP and INITIAL LIQUID. The inhibitory link temporarily reduces the priming of INITIAL LIQUID relative to INITIAL STOP, and the latter is activated with the first pulse from the phonological time node. Once activated, INITIAL STOP strongly primes the entire domain of initial stop nodes and one of these, *p* (initial stop), having just been primed, has greatest priming and becomes activated under the most-primed-wins principle.

Following activation, INITIAL STOP becomes self-inhibited. This releases the inhibition on INITIAL LIQUID, which now achieves the most priming in the domain of phonologic sequence nodes and becomes activated with the next pulse from phonologic time. INITIAL LIQUID therefore strongly primes its domain of nodes, but having just been primed, *r* (initial liquid) achieves greatest priming, reaches threshold soonest, and becomes activated under the most-primed-wins principle.

ERRORS WITHIN THE THEORY

Error-free output occurs under this theory when an "intended-to-be-activated" content node has greater priming than any other node in its domain when the triggering mechanism is applied, that is, whenever the sequence node for the domain of content nodes is activated. The "intended-to-be-activated" node is the one that is receiving priming from a superordinate node in the output sequence, that is, the directly connected content node immediately higher in the hierarchy. This priming summates over time and eventually exceeds the priming of all other nodes in the domain, by time t_1 in Figure 13-3. At this point or any point in time thereafter, the triggering mechanism will activate the intended-to-be-activated node under the most-primed-wins principle and the output is error-free.

Errors occur whenever another node in the domain has greater priming than the intended-to-be-activated node when the triggering mechanism is applied. The fundamental cause of errors is that other, extraneous sources contribute priming which sometimes can exceed the systematically increasing priming for the intended-to-be-activated node when the triggering mechanism is applied. As a consequence, the wrong node becomes activated under the most-primed-wins principle, and an error occurs. Because of the shape of the priming function (see Figure 13-3), errors will be more likely the faster the rate of speech (i.e., the sooner the triggering mechanism is activated following onset of priming (t_0) for every node in the hierarchy).

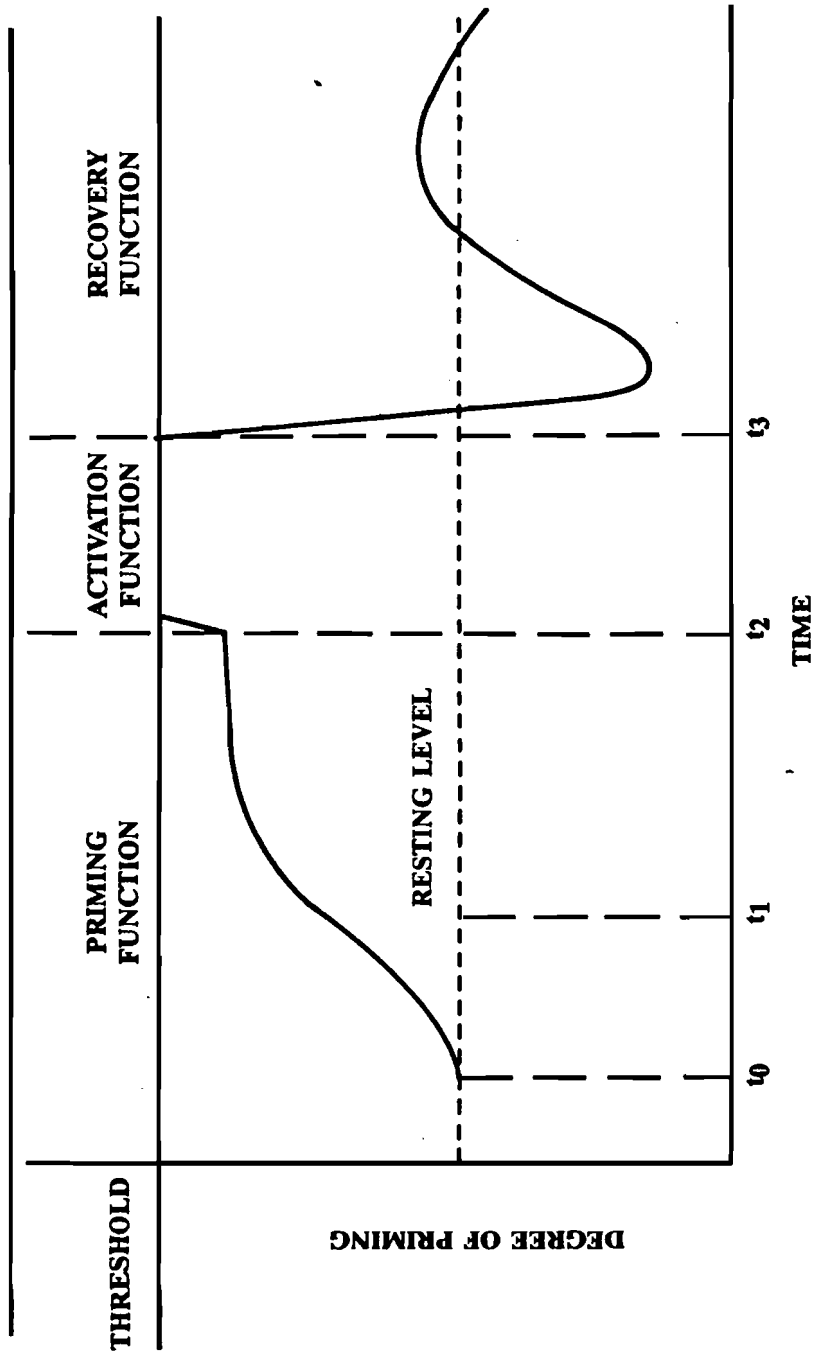


FIGURE 13-3. The priming, activation, and recovery functions for a node within a hypothetical domain of nodes. Onset of priming begins at a time t_0, onset of activation at time t_2, and onset of recovery at time t_3. See text for explanation.

Even Freudian slips are explainable in this way. Consider for example the Freudian substitution of *bottle scarred* for *battle scarred*, spoken of a general who is covertly believed to be incompetent as a result of "hitting the bottle." This covert belief independently primes the node for *bottle*, so that *bottle* (noun) has greater priming than the intended-to-be-activated node, *battle* (noun), when the triggering mechanism is applied to the domain of noun nodes. The wrong node is activated, with the resulting word substitution, because the triggering mechanism automatically activates the most primed node in its domain.

This example, of course, deals with high-level components, but errors involving lower level, phonological, and muscle movement components require a similar explanation. For all substitution errors the theory predicts that the intended and substituted components will belong to the same domain or sequential class (represented in brackets in Figure 13-1). The theory also predicts that, within limits, errors will increase as a function of rate of utterance (see MacKay, 1982).

FEEDBACK WITHIN THE THEORY

Auditory feedback is processed, according to the theory, in the same way as other speech inputs. Sensory analysis nodes constitute the first stage of processing, followed by "mental nodes," the systems of phonological and sentential nodes discussed above. This means that the same mental nodes provide the basis for both producing and perceiving cognitive units such as phrases, words, syllables, and phonemes. Top-down connections between nodes for a word are responsible for producing the word, whereas bottom-up connections are responsible for perceiving it.

Because mental nodes have both bottom-up and top-down connections, the possibility of reverberatory effects that can lead to stuttering exists at every level in every system. The reasoning can be illustrated for two hierarchically connected but otherwise arbitrary nodes: *A* (superordinate) and *B* (subordinate). During production, *A* becomes activated and primes *B* via the top-down connection. However, subsequent activation of *B* could lead to a reactivation of *A* because of the bottom-up connection required for perception involving both these same nodes. Self-inhibition following activation is hypothesized, therefore, to ensure that bottom-up priming resulting from the activation of subordinate nodes does not lead to reactivation of higher level nodes. Following self-inhibition, a normal recovery cycle with rebound from inhibition is posited in which self-priming rises above and then slowly returns to resting level (see Figure 13-3).

Self-inhibition following activation also prevents normal (undelayed) auditory feedback from causing stuttering. Feedback processed by sensory analysis nodes normally returns to and primes the same low-level mental nodes that were responsible for generating the just completed speech output. However, inputs arriving during the period of self-inhibition do not add enough priming to make self-inhibited nodes the most primed in their domain. As a consequence, just activated nodes will not become reactivated when the triggering mechanism is applied to their domain during ongoing production of the remainder of the word or sentence.

INTEGRATION OF STUTTERING INTO THE THEORY

Stuttering is similar to other errors in speech in several respects. Like all other errors, stuttering decreases with repetition or practice in producing a sentence (the adaptation effect; see Brenner, Perkins, & Soderberg, 1972) and with reduction in the rate of speech. Both these effects are readily explained under the theory outlined above (see MacKay, 1982). However, stuttering differs from other speech errors in at least three respects discussed below, which give clues to its etiology within the theory.

The surface characteristics of stuttering

Stutterers exhibit three characteristic phenomena (repetitions, prolongations, and blocks) which differ markedly from other errors in speech. Moreover, these three phenomena predominate in stuttering: Stutterers make other types of errors with no greater frequency than normal speakers.

Repetitions. Repetitions in stuttering usually involve a single consonant or consonant cluster, and only occasionally a syllable or monosyllabic word (Van Riper, 1982). In the present theoretical framework, it is as if nodes, once activated, have a tendency to be reactivated. The reason for this tendency is currently unknown but several possibilities suggest themselves from the theory. One is that the nodes of stutterers may manifest an abnormal priming and recovery cycle (illustrated in Figure 13-4). Under this hypothesis, priming summates abnormally slowly in stutterers, and rebounds abnormally sharply following self-inhibition. As a consequence, a just activated sequence node would have a high probability of being the most primed node in its domain, so that it becomes reactivated with the next pulse from the timing node. The result is, of course, a repetition such as *p-practice*, and the nodes can undergo this cycle again, resulting in a third *p*. However, the cycle cannot go on indefinitely, because reactivated nodes

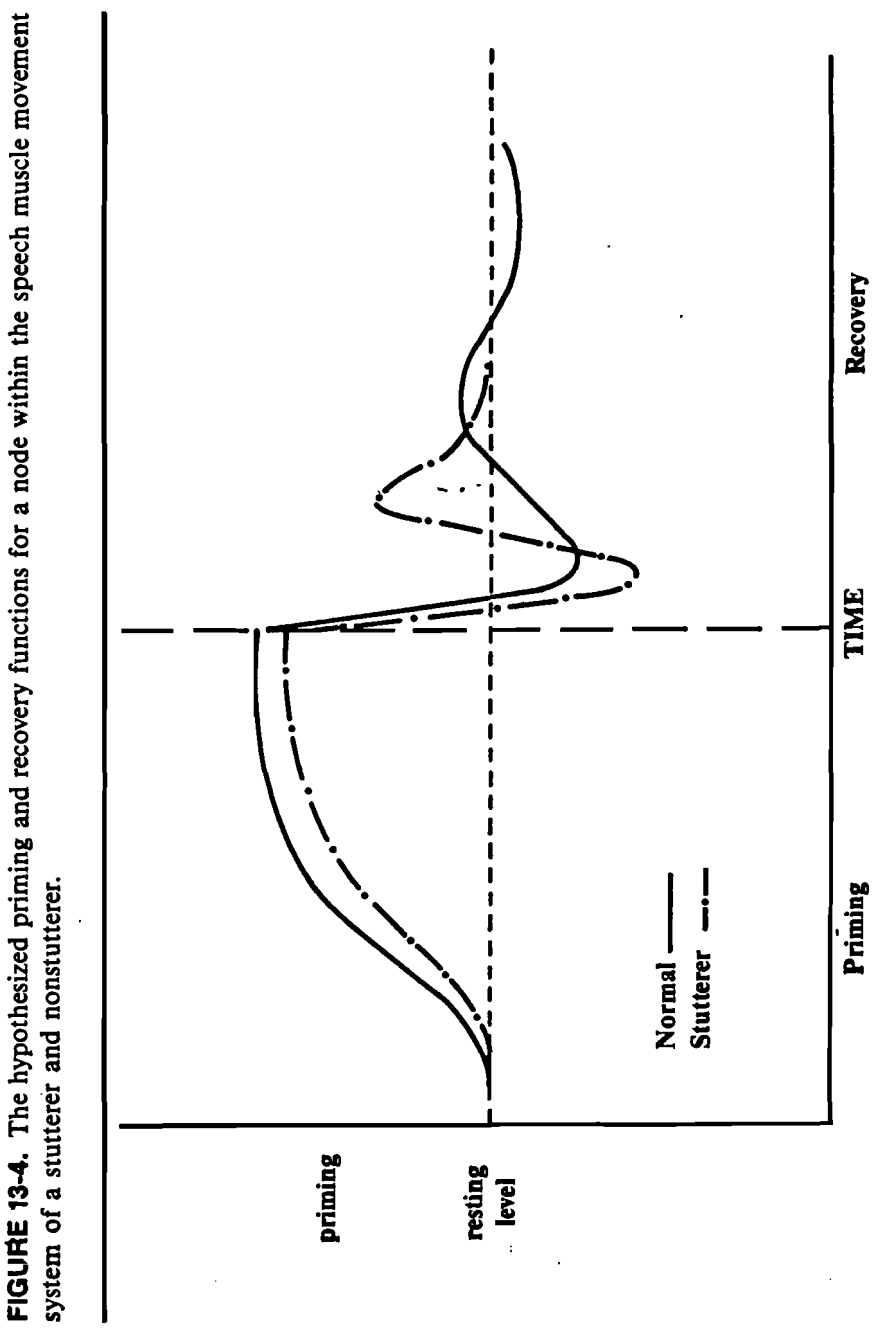


FIGURE 13-4. The hypothesized priming and recovery functions for a node within the speech muscle movement system of a stutterer and nonstutterer.

become fatigued and cannot rise to such high levels of priming on rebound from inhibition. Consequently they no longer achieve more priming than the next-to-be-activated node. Another reason why repetitions eventually stop is that priming for the next-to-be-activated nodes continues to summate during the time that malfunctioning nodes are being reactivated. Thus, the longer the period of stuttering, the more likely that the correct node will be activated with the next pulse from the triggering mechanism.

Major disfluencies occur under this account when several nodes malfunction as a group. However, minor disfluencies, undetectable by the human ear, may occur when only a few muscle movement nodes malfunction. This would explain why stutterers display slower transitions between sounds and greater asynchrony between lip and jaw movements than nonstutterers uttering the same syllables, and why utterances judged by ear to be fluent reveal abnormalities when analyzed by cine-radiography (Zimmerman, 1980).

Prolongations. Prolongations involve the unbroken lengthening of a (continuant) phoneme. A stutterer making this sort of error on *practice* might prolong the *r* to three or four times its normal duration. Descriptively, it is as if the articulators become locked in position during the production of the continuant sound.

Blocks. The most severe problem of stuttering is the inability to utter any sound at all, occurring most often at the beginning of an utterance but also at the beginning of words within an utterance. As Van Riper (1982) pointed out, blocks can be viewed as a special type of prolongation where one or more articulators (the velum, lips, or glottis) are "locked" in an obstructive position, virtually prohibiting airflow and preventing speech.

The cause of blocks and prolongations is currently unknown but the present theoretical framework suggests several possible mechanisms. One is that the nodes of stutterers sometimes fail to terminate activation because of a malfunction in the self-inhibition mechanism. It seems reasonable to suppose that malfunctions in the normal self-inhibition process may be related to the abnormal recovery cycle discussed above (see Figure 13-4), but details of this relationship remain to be worked out.

The distributional characteristics of stuttering

Stuttering occurs mainly at the beginning of words and utterances. The reason, under the theory, can be attributed to "anticipatory priming." Note the numbers in Figure 13-1, which indicate the order in which the corresponding nodes must be activated for error-free output. Activating Node 1 simultaneously primes Nodes 2 and 8, but since 8 can only be activated after 2, 3, 4, 5, 6, and 7 have been activated, the priming of 8 constitutes "anticipatory priming" which summates during the interval that these other nodes are being activated. Anticipatory priming may, therefore, reduce the probability of stuttering by overcoming the slow buildup of priming in the intended-to-be-activated nodes of stutterers for the later components of a word or sentence (see Figure 13-4).

The level at which stuttering originates

Some stuttering theories attribute the problem to the highest level motivational and belief systems of stutterers (Sheehan, 1958; Johnson, 1938, respectively), well above the highest level (sentential) system discussed here. Others attribute the problem to lower levels, for example, the muscle movement system (Perkins et al., 1979), or even sensory systems (e.g., the stapedial reflex at the ear drum, Webster & Lubker, 1968). Still others such as Mysak (1960) contend that stuttering occurs at several levels and involves output components of many different sizes. However, there are four major reasons for believing that stuttering can be localized within the muscle movement system.

First, stutterers do not report stuttering during internal speech (J. Sheehan, personal communication).[1] Since stutterers frequently report stuttering when speaking aloud to themselves (Van Riper, 1982), reduced speaking anxiety associated with interpersonal communication cannot fully account for the absence of stuttering during internal speech. By way of contrast, other speech errors have been observed to occur with equal frequency in internal and overt speech. For example, Dell (1980) found that spoonerisms (e.g., *throat cutting* misproduced as *coat thrutting*) were equally frequent when normal subjects produced tongue twisters either aloud or to themselves. Within the present theory, this finding suggests that spoonerisms, unlike stuttering, can be localized within the phonological system rather than the muscle movement system. Secondly, the probability of stuttering increases with the number of muscle movement components that are involved. Stutterers are most fluent when they are instructed only to move their lips, less fluent when whispering, and least fluent when engaging in full-fledged articulation (Brenner et al., 1972). These findings are clearly consistent with the hypothesis that stuttering involves a disorder within the muscle movement system. Third, stutterers appear to have difficulty with the muscle movements for speech, independent of the phonological system which normally controls the overall sequencing and timing of these movements. For example, in response to a pure tone stimulus, some stutterers are slower than nonstutterers in initiating lip closure

(McFarlane & Prins, 1978) and laryngeal voicing (Cross & Luper, 1979) but not throat clearing or finger pressing (Reich, Till, & Goldsmith, 1981). Finally, stutterers show no deficits in the perception of speech. They never misperceive someone to say "pppplease" when in fact the person said "please." This fits the hypothesis that stuttering does not originate in the sensory analysis nodes (which are specific to perception) or in the phonologic nodes (which govern both perception and production) but rather in the muscle movement nodes (which are specific to production).

The hypothesis that stuttering reflects a malfunction within the muscle movement system, of course, does not imply that higher level processes cannot contribute to the frequency of stuttering. High-level factors such as anxiety and syntactic ambiguity can affect motor control at every level and thereby influence the probability of stuttering (see MacKay, 1969a).

STUTTERING AND THE PROCESSING OF AUDITORY INPUT

So far we have been viewing stuttering as a disorder in the control of sequencing and timing. To determine whether this view is compatible with data relating stuttering to audition, we now review five observations which point to a connection between stuttering and the processing of auditory input: (1) Stuttering can be virtually eliminated for some stutterers with the flick of a switch introducing white noise within the frequency range of speech and loud enough to mask the stutterer's returning auditory feedback (Shane, 1955). (2) The stapedial reflex of the middle ear appears to differ between stutterers and nonstutterers. The stapedius muscle normally contracts 100–165 msec prior to phonation, thereby reducing the amplitude of eardrum vibration and attenuating the hearing of one's own voice. Webster and Lubker (1968) found that the stapedial reflex is less stable in stutterers than nonstutterers, and Horovitz, Johnson, Pearlman, Schaffer, and Hedin (1978) found that under conditions of anxiety, stutterers show less stapedial attenuation than nonstutterers. (3) Auditory processing of an about-to-be-produced word appears to guide and facilitate its production. For example, stutterers often release themselves from a block when someone else utters the word on which they are having difficulty (Barr & Carmel, 1970). Similarly, stuttering is ameliorated when stutterers shadow words they hear or produce words in unison with another speaker. (4) Some stutterers become more fluent when their returning auditory feedback is delayed by means of a recording and reproducing device (Huchinson & Burke, 1973; Novak, 1978; Preus, 1981; Webster, Shumacher, & Lubker, 1970). (5) Repetition errors resembling

those of stutterers can be obtained in normal individuals by amplifying as well as delaying their returning auditory feedback. The effects of delayed auditory feedback have been the object of a large number of studies which we summarize briefly below. For adults, repetition errors due to delayed auditory feedback increase as a function of delay up to about .2 sec, and then decrease, with longer delays, but never disappear completely even with delays as long as .8 sec (MacKay, 1968). The peak delay producing the greatest number of repetition errors lengthens as an inverse function of age; .2 sec for adults, .375 for children aged 7–9, and approximately .75 sec for children aged 4–6 (MacKay, 1968), but not as a function of language familiarity: When bilinguals speak either their more or their less familiar language under delayed auditory feedback, their peak delay remains the same (MacKay, 1969b). Language familiarity only influences the *degree* of disruption at any given delay. Bilinguals make more repetition errors when producing their less familiar language under delayed auditory feedback (MacKay, 1970b), but practice in producing a sentence reduces the number of errors when subjects subsequently produce the sentence under delayed auditory feedback (MacKay, 1970b). The amount of disruption also diminishes with age, since practice, familiarity, and experience in producing speech increase as children grow older (MacKay, 1968). Mechanical distortions of the returning auditory feedback likewise reduce the disruptions resulting from delayed auditory feedback (Hull, 1952; see also MacKay, 1969c).

Theoretical explanation of the auditory effects

Until quite recently, feedback control theory has provided the main framework for explaining the relation between stuttering and the processing of auditory feedback (see Mysak, 1960). We begin, therefore, by outlining the nature of feedback control theory and its problems before addressing the auditory effects within the present theory. Under feedback control theory, feedback from sensory systems plays a direct role in controlling ongoing action, so that delaying this feedback results in control errors such as stuttering. Feedback control theory has never achieved a detailed explanation of instances of stuttering (Garber & Seigel, 1982), and on close inspection fails to explain many of the general effects as well. Consider delayed auditory feedback, for example. Why is there a delay that produces maximal disruption of speech? Under feedback control theory, disruption should either remain constant or increase monotonically as a function of delay. Why is it necessary to amplify the returning feedback in order to bring about articulatory errors? Why do subjects speak louder when their amplified auditory feedback is delayed? Under feedback control theory, amplified feedback

is that it postulates a specific hypothesis as to the level in the speech production system that stuttering arises. The theory also integrated two major approaches to stuttering which have been developing independently over the past several decades, one considering stuttering as a disorder of motor control and the other as a disorder in the processing of auditory feedback (see Garber & Seigel, 1982).

Futher research is needed to test the theory proposed here. More importantly, further theoretical work within the framework of the metatheory proposed here is needed to bridge the long-standing gap between normative psychology and stuttering research. Recent years have seen a great deal of research into errors in behavior (see for example Fromkin, 1980), and this area is likely to become a major concern of psychological models over the next decade (see Neisser, 1982). It would be unfortunate if stuttering does not become part of that larger concern since bridging the gap has benefits for both psychology and stuttering research. Under the metatheory proposed here, an understanding of the processes underlying fluent speech is necessary in order to understand its disruption in a complex speech disorder such as stuttering. But it is equally true that transient malfunctions such as stuttering need to be explained in theories of normal behavior, since a complete and adequate theory must be capable of predicting all of the ways that an output system will break down.

NOTES

[1] C. Van Riper, in a subsequent personal communication, reports that some stutterers do claim to stutter during internal speech. This unexpected difference in expert opinion suggests that this issue warrants empirical investigation in the manner illustrated in MacKay (1982, Figure 1) and Dell (1980).

[2] For purposes of exposition this explanation has been simplified by ignoring (a) the time it takes for the muscles to come into action following onset of activation of the muscle movement nodes, (b) the time it takes airborn auditory feedback to arrive at the ears, (c) the time it takes the sensory analysis nodes to process the feedback and delivery bottom-up priming to the lowest level phonologic nodes, and (d) the time it takes to pass this priming on to the just activated muscle movement nodes.

REFERENCES

Adams, J. A. Issues for a closed-loop theory of motor learning. In G. E. Stelmach (Ed.), *Motor control: Issues and trends.* New York: Academic Press, 1976.

Barr, D. F., & Carmel, N. R. Stuttering inhibition with the high frequency narrow band masking noise. *Journal of Auditory Research,* 1970, *10,* 59-61.

Bloodstein, O. *Handbook of stuttering.* Chicago: National Easter Seal Society for Crippled Children and Adults, 1969.

Brenner, N. C., Perkins, W. H., & Soderberg, G. A. The effect of rehearsal on frequency of stuttering. *Journal of Speech and Hearing Research,* 1972, *15,* 483-486.

Cooper, M. H. & Allen, G. D. Timing control accuracy in normal speakers and stutterers. *Journal of Speech and Hearing Research,* 1977, *20,* 55-71.

Cross, D. E., & Luper, H. L. Voice reaction time of stuttering and nonstuttering children. *Journal of Fluency Disorders,* 1979, *4,* 661-68.

Dell, G. *Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally elicited speech errors.* Unpublished doctoral dissertation, University of Toronto, 1980.

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

Garber, S. R., & Siegel, G. M. Feedback and motor control in stuttering. In D. K. Routh (Ed.), *Learning, speech, and the complex effects of punishment.* New York: Plenum Press, 1982.

Horovitz, L. J., Johnson, S. B., Pearlman, R. C., Schaffer, E. J., & Hedin, A. K. Stapedial reflex and anxiety in fluent and disfluent speakers. *Journal of Speech and Hearing Research,* 1978, *21,* 762-767.

Huchinson, J. M., & Burke, K. M. An investigation of the effects of temporal alteration in auditory feedback upon stutterers and clutterers. *Journal of Communication Disorders,* 1973, *6,* 193-205.

Hull, F. M. *Experimental investigation of speech disturbance as a function of the frequency distortion of delayed auditory feedback.* Unpublished PhD thesis, University of Illinois, 1952.

Johnson, W. The role of evaluation in stuttering behavior. *Journal of Speech Disorders,* 1938, *3,* 85-89.

MacKay, D. G. Metamorphosis of a critical interval: Age-linked changes in the delay in auditory feedback that produces maximal disruption of speech. *Journal of the Acoustical Society of America,* 1968, *19,* 811-21.

MacKay, D. G. Effects of ambiguity on stuttering: Towards a theory of speech production at the semantic level. *Kybernetik,* 1969, *5,* 195-208. (a)

MacKay, D. G. Forward and backward masking in motor systems. *Kybernetik,* 1969, *6,* 57-64. (b)

MacKay, D. G. To speak with an accent: Effects of nasal distortion on stuttering under delayed auditory feedback. *Perception and Psychophysics,* 1969, *5,* 183-88. (c)

MacKay, D. G. Context-dependent stuttering. *Kybernetik,* 1970, *7,* 1-9. (a)

MacKay, D. G. How does language familiarity influence stuttering under delayed auditory feedback? *Perceptual and Motor Skills,* 1970, *30,* 655-69. (b)

MacKay, D. G. The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology,* 1972, *3,* 210-27.

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

McFarlane, S. J., & Prins, D. Neural response time of stutterers and nonstutterers. *Journal of Speech and Hearing Research,* 1978, *21,* 768-78.

Mysak, E. Servo theory and stuttering. *Journal of Speech and Hearing Disorders,* 1960, *25,* 188-195.

Novak, A. The influence of delayed auditory feedback on stutterers. *Folia Phoniatrica,* 1978, *30,* 278-85.

Neisser, U. Understanding psychological man. *Psychology Today,* 1982, *16,* 40.

Perkins, W. H., Bell, J., Johnson, L., & Stocks, J. Phone rate and the effective planning time hypothesis of stuttering. *Journal of Speech and Hearing Research,* 1979, *22,* 747-55.

Preus, A. *Identifying subgroups of stutterers.* Oslo: University of Oslo Press, 1981.

Reich, A., Till, J., & Goldsmith, H. Laryngeal and manual reaction times of stuttering and nonstuttering adults. *Journal of Speech and Hearing Research*, 1981, *24*, 192–196.

Shane, M. L. S. Effect on stuttering of alteration of auditory feedback. In W. Johnson (Ed.), *Stuttering in children and adults*. Minneapolis: University of Minnesota Press, 1955.

Shane, M. L. S. Effect on stuttering of alternation in auditory feedback. In W. Johnson & R. R. Levitenneger (Eds.), *Stuttering in children and adults*. Minneapolis: University of Minnesota Press, 1955.

Sheehan, J. G. Conflict theory of stuttering. In J. Eisenson (Ed.), *Stuttering: A symposium*. New York: Harper & Row, 1958.

Siegel, G. M., & Pick, H. L. Auditory feedback in the regulation of voice. *Journal of the Acoustical Society of America*, 1974, *56*, 1618–1624.

Treiman, R., Salasoo, A., Slowiaczek, L., & Pisoni, D. Effects of syllable structure on adults' phoneme monitoring performance. *Research on Speech Perception Progress Report* 8. Bloomington: Indiana University, 1982.

Van Riper, C. *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

Webster, R., & Lubker, R. B. Interrelationships among fluency producing variables in stuttered speech. *Journal of Speech and Hearing Research*, 1968, *11*, 219–23.

Webster, R. L., Schumacher, S. J., & Lubker, R. B. Changes in stuttering frequency as a function of various intervals of delayed auditory feedback. *Journal of Abnormal Psychology*, 1970, *75*, 45–49.

Zimmerman, G. N. Articulatory behavior associated with stuttering: A cineradiographic analysis. *Journal of Speech and Hearing Disorders*, 1980, *23*, 108–21.