KYBERNETIK

BAND VII APRIL 1970 HEFT 1

Context-Dependent Stuttering

Donald G. MacKay*
Department of Psychology, University of California, Los Angeles, California 90024

Received October 26, 1969

Abstract. Bei einer Untersuchung des Stotterns in der deutschen Umgangssprache machten wir folgende Beobachtungen:

1. Das gestotterte Phonem wird häufig von einem identischen Phonem begleitet (definiert als das „induzierende Phonem") welches sowohl vor wie nach dem gestotterten Phonem auftreten kann.
2. Gewöhnlich folgte das induzierende Phonem dem gestotterten Phonem.
3. Der Abstand zwischen induzierendem und gestottertem Phonem war geringer als bei Zufälligkeit zu erwarten.
4. Induzierende und gestotterte Phoneme befanden sich gewöhnlich in identischen Silbenpositionen.
5. Gestotterte Phoneme traten in der Regel bei betonten Silben auf.

Um diese Beobachtungen zu erklären erschienen uns drei Annahmen erforderlich:

1. Sprach-Output ist hierarchisch bestimmt. Silben und Phoneme sind Glieder in dieser Hierarchie.
2. Unterschwellige Erregbarkeit ist in dieser Hierarchie stärker bei betonten als bei unbetonten motorischen Programmen.
3. Ähnliche Programme (sowohl auf Silben- als auch Phonemniveau) inhibieren einander.

Diese Annahme gibt zugleich eine mögliche Erklärung für Blockierung und Längung — Phänomene, die ebenfalls in der Sprache von Stotterern auftreten. Unsere Beobachtungen bieten also eine mögliche Lösung für das Rätsel des Stotterns.

A solution to the riddle of stuttering has both practical and theoretical significance. The theoretical significance of stuttering lies in the fact that transient malfunctions in motor output are crucial facts to be explained in theories of behaviour in the same sense that illusions are crucial for theories of perception (Teuber, 1960). An adequate model of a motor system must be capable of predicting how and when the motor system will break down. Yet despite the practical and theoretical importance of solving the riddle of stuttering, the actual cause of stuttering remains as much a mystery today as when it was first systematically studied in 1868 (Wyneken, 1868).

The term stuttering in the American literature usually embraces three distinct behaviours: involuntary repetition of a speech sound, prolongation of a speech sound and blocking or transitory inability to execute a speech sound. But since an explanatory model must ultimately be capable of explaining these various final outputs, I will use the term stuttering to refer exclusively to involuntary repetition of speech sounds, and treat prolongation and blocking separately. However, the model of stuttering that I have developed also predicts blocking and prolongation of speech sounds in certain circumstances, which is desirable for a general model of this disturbance, since prolongation and blocking frequently appear alongside stuttering in the natural speech of stutterers.

This paper deals with the phonetic context of stuttered phonemes. A large number of factors have been shown to influence stuttering in the thousands of papers on the subject, but the phonetic context of stuttering has been little studied. This oversight is surprising since it has long been known that phonetic context directly contributes to other errors in speech. For example, omissions of phonemes can usually be attributed to the prior or subsequent occurrence of an identical phoneme in the speech context (MacKay, 1969).

Thus the first L in *Unglaublich* can be said to induce the omission of the second L.

1. Unglaublich → Unglaubich[1] [from Meringer and Mayer, 1895].

My research indicates a similar contextual factor in the stuttering of normal individuals (with no history of chronic stuttering). My approach resembled that of neurologists such as John Hughlings Jackson and ethologists such as N. Tinbergen in attempting to infer the properties of a complex mental system from the nature of its output, rather than the nature of its input. And in so far as stuttering is not a random disturbance of the motor system, such inferences are of interest to linguists, psychologists, and neurophysiologists as well as speech pathologists.

My data consisted of one hundred and twenty-six instances of stuttering in conversational speech. The raw data, the stutters and the sentential contexts in which they occurred, were compiled and published in German in 1895 and 1908 by Rudolf Meringer. Meringer's procedures for data collection have much to recommend them and the reader is referred to MacKay (1969) for a detailed discussion of these methods.

The basic question in the present study was whether contextual factors such as recurring phonemes in a sentence contribute to the occurrence of stuttering. For each contextual factor I examined, I tested a null hypothesis based on the structure of normal, nonstuttered speech. Of course I wanted the nonstuttered speech to be representative of Meringer's speakers, most of whom were professors at the University of Vienna. But since Meringer only recorded the errors

* This work was supported in part by an MIT intramural fellowship, a NASA grant (NsG 496) to Dr. H. L. Teuber, UCLA grant 2428, and USPHS Grant HM-16, 668. The author thanks H. Schulze for his help in translating Meringer (1908) and Meringer and Mayer (1895).

1 The arrow in these formulae is synonymous with "was spoken as".

his speakers made, I used a corrected set of sentences containing word reversals as a corpus of nonstuttered speech.

Recurring Phonemes

(1) The Contextual Induction Hypothesis

The Contextual Induction Hypothesis predicts that stuttering on a phoneme can be induced by an identical phoneme earlier or later in the context of the sentence. Tests of this hypothesis rest on whether stuttered phonemes are preceded or followed by an identical phoneme more often than would be expected by chance. In our test of this hypothesis an instance of contextual induction was counted, whenever the recurring phoneme occured in either an immediately adjacent word or the same word as the stuttered phoneme. Now, under the null hypothesis the frequency of phoneme repetition in the same or immediately adjacent words in the corpus of natural speech should be the same as in the corpus of stuttered speech.[2] These frequencies were calculated, with the results shown in Table 1.

Table 1. *The percent of the corpus (stutters in natural speech) involving repeated phonemes. Chance is calculated from the frequency of phoneme repetition in natural speech*

	Repeated phonemes	No repeated phonemes
Data	83	17
Chance	14	86

As can be seen there, an identical phoneme preceded or followed a stuttered phoneme much more frequently than would be expected under the null hypothesis.[3] A stuttered phoneme preceded or followed an identical phoneme in the context more often than would be expected by chance, a result statistically significant at the 0.01 level using a Chi Square test. Similar reliability was obtained when the probability of phoneme repetition within a single word was calculated and compared to chance expectation.

Thus our data indicate a correlation between stuttering and the earlier or later occurrence of an identical phoneme in the context. In principle, of course, a correlation between two factors, A and B, could be caused by some third factor, that effects both A and B, or A could have caused B, or B could have caused A. But only one of these alternatives seems reasonable here since stuttering cannot possibly influence the phonemes that compose a word. Nor is it likely that some third factor causes both stuttering and the phoneme structure of words in which stuttering occurs. Rather the identical phoneme which precedes or follows a stuttered phoneme must somehow induce or contribute to the stuttering. Thus the Contextual Induction Hypothesis seems to provide the only reasonable interpretation of our data.

2 The exact mathematical formulae for calculating this and other null hypotheses are presented in MacKay (1969).

3 Sometimes more than one word separated the stuttered and inducing phonemes. For example, the *ein* in *einmal* may have induced the stuttering in the following sequence:

einmal das Recht ein-einräumt (from Meringer).

But since more than one word separated *einmal* and *einräumt*, this was not counted as an example of contextual induction in our analysis.

(2) Anticipation Hypothesis

One explanation of the Contextual Induction phenomenon is that stuttering results from anticipation of an identical phoneme to come. Hocket (1965), outlines this Anticipation Hypothesis in detail, suggesting that the inducing phoneme should more frequently follow than precede the stuttered phoneme. Cases where the inducing phoneme followed the stuttered phoneme were defined as Proactive Induction and cases where the inducing phoneme preceded the stuttered phoneme were defined as Retroactive Induction.

2. Fredrick → Fr-riedrick.
3. muss man → m-muss man.
4. Stuck steht → Stuck st-steht.
5. die Details → die D-Details.

Examples of Retroactive Induction are presented in 2 and 3 (inducing phoneme underlined), and Proactive Induction in 4 and 5.

Our data showed that Proactive Induction was much more frequent than Retroactive induction as predicted in the Anticipation Hypothesis. The inducing phoneme followed the stutter in 74% of the cases, and preceded the stutter in only 26% of the cases of phoneme induction (see Table 2). A Chi Square test showed that Proactive Induction was significantly more frequent than Retroactive Induction (0.05 level) confirming the prediction of the Anticipation Hypothesis.

Table 2. *Stuttering as a function of whether the inducing phoneme preceded or followed the stuttered phoneme. The data are in percent*

	Proactive induction	Retroactive induction
Data	74	26
Chance	50	50

However, this finding does not preclude the reality of Retroactive Induction. To reach this conclusion we must show that Retroactive Induction occurs no more frequently than would be expected by chance. The data are shown in Table 3. As can be seen there, both

Table 3. *The repeated phoneme effect: a comparison of backward and forward induction*

	Repeated phonemes		No repeated phonemes
	Proactive induction	Retroactive induction	
Data	61	22	17
Chance	7	7	86

Proactive and Retroactive Induction occurred with greater than chance frequency. So Retroactive Induction, although less frequent than Proactive Induction is nevertheless a real phenomenon. By implication then, the Anticipation Hypothesis cannot serve as a general explanation of context-dependent stuttering.

(3) Interaction Hypothesis

One explanation of context-dependent stuttering (outlined in the discussion) postulates interaction between similar phonemes in close proximity in the serial order of speech. According to this hypothesis, inducing phonemes should occur closer to stuttered phonemes than would be expected by chance. In testing this hypothesis only cases where the stuttered and inducing phoneme occurred in the same word were examined, so as to facilitate construction of the null hypothesis. Under the null hypothesis stuttered and inducing phonemes should occur no closer together than repeated phonemes in the corpus of natural speech.

We measured the proximity of stuttered and inducing phonemes in both phonemes and number of syllables, but since both analyses gave similar results, only the syllable analysis will be discussed here. Since stuttered and inducing phonemes always occurred in different syllables, we represented an instance of stuttered and inducing phonemes in immediately adjacent syllables as separation 0; one intervening syllable was separation 1; and so on.

The results of this analysis are shown in Fig. 1, where it can be seen that stuttered and inducing

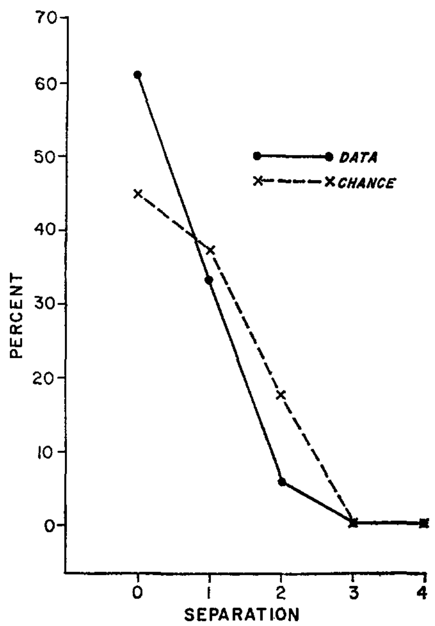

Fig. 1. The proximity (in syllables) of stuttered and inducing phonemes (solid line). The broken line represents the chance proximity, based on the separation of repeated phonemes in non-stuttered speech of these speakers

phonemes occurred closer together than would be expected by chance. Thus the prediction of the Interaction Hypothesis was confirmed.

A second point of interest is that the proximity analysis showed no difference between Proactive and Retroactive Induction. The stuttered and inducing phonemes occurred closer than chance expectation for both Proactive and Retroactive Induction. This finding suggests again that a general model of stuttering must account for both Proactive and Retroactive Induction as well as the proximity of the stuttered and inducing phonemes.

(4) Allophonic Similarity Hypothesis

Phonemes in natural speech are modified to fit the context in which they occur, the contextual variants of phonemes being termed allophones. And although the native speaker usually perceives the various allophones of a phoneme as the same, each allophone is physiologically distinct. The reader can readily demonstrate this for himself. Hold a lighted match in front of the lips and say the words SPIN and PIN at normal conversational loudness. The P in PIN will blow out the match but the P in SPIN will not, indicating a physiological difference between these allophones of P.

Our question here was whether the stuttered and inducing phonemes are also indentical as allophones. This question reduces to whether the stuttered and inducing phonemes are not only identical as phonemes, but occur in indentical phonetic contexts.For example in 9 identical phonemes precede and follow the stuttered and inducing phoneme, representing a case of identity of the stuttered and inducing allophones.

6. machmal → m-manchmal.

But these contextual phonemes were identical less frequently than chance expectation where chance is the probability that phonemes immediately preceding or following repeated phonemes were identical in the corpus of natural speech.

Now the similarity of two phonemes is reflected in the number of "distinctive features" they have in common. Phonemes sharing all but one distinctive feature are most similar, those sharing all but two are less similar and so on. We used the standard ICL distinctive feature system (as modified by Wickelgren, 1966) in our analysis, since it was based on articulatory rather than perceptual criteria (which seemed less relevant to errors in speech production). This feature system contains four underlying features, so that a 4 was scored when the contextual phonemes were indentical (as in 6) a 3 when all but one feature was identical as in 7 where r and l (the phonemes following the stuttered and inducing phonemes) differ in only one feature.

7. Schrift schliesst → Schrift sch-schliesst
   (phonetic spelling after Meringer).

The average similarity of the contextual phonemes is shown in Fig. 2, where it can be seen that phonemes preceding (or following) stuttered and inducing phonemes differed in one or two distinctive features more frequently than would be expected by chance. But these contextual phonemes were less frequently identical or completely different than chance expectation. This suggests that the stuttered allophones differed from their inducing allophones. Thus the final motor units for stuttered and inducing phonemes must be highly similar but not completely identical.

(1.1) Syllabic Factors

The Syllabic Similarity Hypothesis

The Syllabic Similarity Hypothesis predicts that stuttered and inducing phonemes will occur in identical syllable positions more often than would be expected by chance. Four syllabic positions of consonants were examined in a test of this hypothesis. The syllabic positions were initial, next to initial, final and next to final. For example in the monosyllabic word STAND, S occurs in initial syllabic position, T in next to initial position, N in next to final and D in final position. Of course in a CVC syllable such as TAN, T occurs in initial position and N in final position.

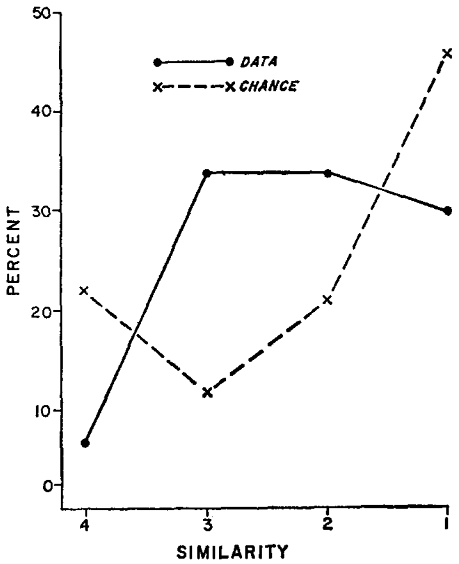

Fig. 2. The similarity or number of identical distinctive features of the phonemes surrounding the stuttered and inducing phonemes

Table 4. *A comparison of the syllabic position of stuttered and inducing phonemes*

	Syllabic position	
	Same	Different
Data	81	19
Chance	29	71

The data are shown in Table 4 where it can be seen that stuttered and inducing phonemes took the same syllabic position more frequently than would be expected by chance, where chance was calculated as the frequency with which identical phonemes had identical syllabic positions in the corpus of natural speech. Thus, the data supported the Syllabic Similarity Hypothesis.

The Syllabic Structure Hypothesis

Further analyses showed that stuttering occurred in initial syllabic position and in the initial syllable of words more often than would be expected by chance, indicating an interesting parallel with the findings of Brown (1936) for pathological stutters. These findings also support the Syllable Structure Hypothesis proposed by Hocket (1965) and MacKay (1969). The Syllable Structure Hypothesis maintains that final consonants are grouped with vowels, and subgroups are broken up with greater difficulty than isolated elements. Phonemic Stuttering is of course an instance where a syllable is broken up. Thus, the Syllable Structure Hypothesis can be viewed as a tentative explanation of why initial consonants, not being part of a syllabic subgroup, are stuttered more frequently than final consonants which are part of the vowel group according to this theory.

(1.1.1) Stress Factors

The Stress Re-entry Hypothesis

The Stress Re-entry Hypothesis predicts that stressed elements will re-enter or be stuttered more often than unstressed elements. To test this hypothesis we examined the stress pattern of the words containing stuttering, determining whether the stuttered phoneme occurred in a stressed or unstressed syllable. The data showed that stuttered phonemes were usually stressed, regardless of whether the inducing phoneme was stressed or unstressed. For example, in 8 and 9 the stuttered phoneme occurs in a stressed syllable even though the inducing phoneme (underlined) occurs in an unstressed syllable, whereas in 10 both the stuttered and inducing phoneme are stressed.

8. Publikum kommt → Publikum k-kommt
9. manchmal → m-manchmal
10. leicht erklärlich → l-leicht erklärlich.

This stress factor was systematically analyzed with the results shown in Table 5 where it can be seen that 84% of the stuttered phonemes occurred in stressed syllables, in comparison to the 34% expected by chance.

Table 5. *The syllabic stress of stuttered phonemes. Chance is calculated from multisyllabic words in natural speech*

	Stuttered phoneme stressed	Stuttered phoneme unstressed
Data	84	16
Chance	34	66

However, although this difference was statistically significant, we wondered whether it might not simply reflect the fact that stuttering occurs most frequently in word-initial syllables which are usually stressed in German. We therefore carried out a separate analysis of stuttering on noninitial syllables. The data are shown in Table 6 where it can be seen that phonemes in noninitial syllables were stuttered more frequently than would be expected by chance. Thus stress must have an effect of its own, independent of syllabic position, supporting the Stress Re-entry Hypothesis.

Table 6. *The syllabic stress of stuttered phonemes in initial and noninitial syllables. Chance is calculated from multisyllabic words in natural speech*

	Noninitial phonemes	
	Stuttered phoneme stressed	Stuttered phoneme unstressed
Data	64	36
Chance	25	75
	Initial phonemes	
	Stuttered phoneme stressed	Stuttered phoneme unstressed
Data	88	12
Chance	66	34

Discussion

Our data indicate that an adequate model of stuttering in the natural speech of normal individuals must explain three main sets of factors: the recurrent phoneme effect (contextual induction); syllabic factors; and stress factors.

The goal of this paper was to develop a model incorporating these factors with the added constraint that the model must be consistent with mechanism underlying other types of errors in speech (see Fig. 3).

In line with this constraint we began by comparing the conditions under which masking (phoneme omission) and stuttering occur.

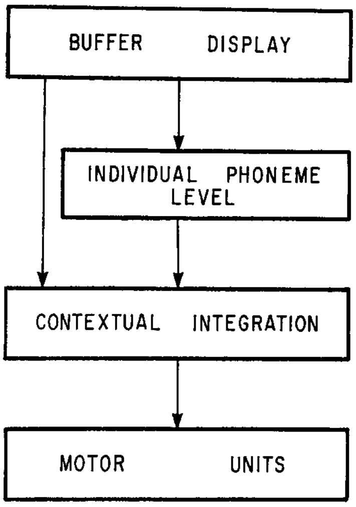

Fig. 3. An oversimplified model of speech production at the phonetic level (based on studies of Spoonerisms and masking). A buffer system displays the phonemes of the words we intend to say, perhaps a phrase at a time. The buffer feeds into two levels: an individual phoneme level containing the internal representatives of phonemes; and a contextual integration level containing programmes for modifying phonemes, based on the context in which they occur. Finally the motor units code the contextual variants of phonemes. When two contextual variants at this level are highly similar, they are assumed to interact in mutually inhibitory fashion. The crucial component for speech production in the model is not represented, namely a scanning device that serially activates the phonemes displayed in the buffer system, thereby determining both the serial order and rate of speech

Comparison of Masking and Stuttering Facts:

Phoneme masking is defined as the omission of one phoneme in a sequence due to the prior or subsequent occurrence of another phoneme in the sequence. The omitted phoneme is termed the masked phoneme and the other the masking phoneme. As pointed out earlier, the masking phoneme is almost invariably identical to the masked phoneme. For example, the first R in FRIEDRICH is the masking phoneme and the second the masked phoneme in 11.

11. Friedrich → Friedich
12. repress → repess (from Bawden, 1900).

Curiously, however, the allophones of the masked and masking phoneme differ more frequently than would be expected by chance. Finally, the masked phonemes usually occur in unstressed rather than stressed syllables.

Theory

Two basic assumptions seemed necessary to explain context dependent masking.

(1) Reciprocal Inhibition

One assumption is that motor units coding similar (but not completely identical) allophones interact in mutually inhibitory fashion. Thus, the motor units coding the allophones for the two L's in *Unglaublich* are assumed to interact in this way before the word is produced.

(2) The Stress Assumption

Subthreshold excitability for stressed elements was assumed to be greater than for unstressed elements. Thus, the summation time for final activation of stressed elements will be less than for unstressed elements. From these assumptions we can explain the fact that unstressed elements will be masked more frequently than stressed elements. Consider the case of backward masking. If A and B are the interacting units for a repeated phoneme in the programme, and if a and b are the summation times for activating A and B, then backward masking will occur when:

13. $a - b > d + b$ or equivalently, $a > d + 2b$

where d is the lag between onset of scanning and A and B. A model of these hypothetical processes is shown in Fig. 4.

Context-independent omission of unstressed units as described in Heffner (1963) can also be explained in this model. Specifically, context-independent omission of A will occur when

14. $E(t) < a$

where t is the increment in excitability of unit A for the period between onset and offset of scanning of A. Again, of course, A will most likely be omitted when unstressed (i.e. when a is large) or when the rate of speech is rapid (i.e. when t is small).

Towards a Model Stuttering at the Phonetic Level

(1) Oscillation

Our present goal is to develop a model of stuttering which is consistent with this model of masking. Consider the assumption that similar allophones interact in mutually inhibitory fashion. One of the properties of such a reciprocal inhibitory arrangement, known as Oscillation, provides a possible explanation of stuttering. Oscillation refers to the fact that when one of the components of a reciprocal inhibitory network becomes phasically activated, the excitability of both components will oscillate.[4] Now, when the excitability of one

4 The physiological basis for oscillation is well understood. Imagine two units A and B having a reciprocal inhibitory relationship. When one of the units (A) is activated, the other component (B) will be inhibited. But once activation of A ceases, B will rebound or undergo a supra-normal phase known as hyperexcitability, becoming more activated than usual (Bullock, 1965). Now, of course, this means that A will become inhibited by the hyperexcitability of B, so that a cycle is set up, with oscillation in excitability of A and B for a period which depends on the damping or refractory properties of these units.

of the units is very close to threshold, the peak in oscillatory excitability may exceed threshold, and each time this happens, stuttering will occur. Thus, the reciprocal inhibitory assumption in masking models may also explain the repeated phoneme effect in context-independent stuttering.[5]

Finally, the question arises as to why repeated phonemes sometimes result in masking and sometimes in stuttering. The answer to this question must lie in part in the programming of stress, since stuttered phonemes are usually stressed, whereas masked phonemes are usually unstressed. This factor is discussed below.

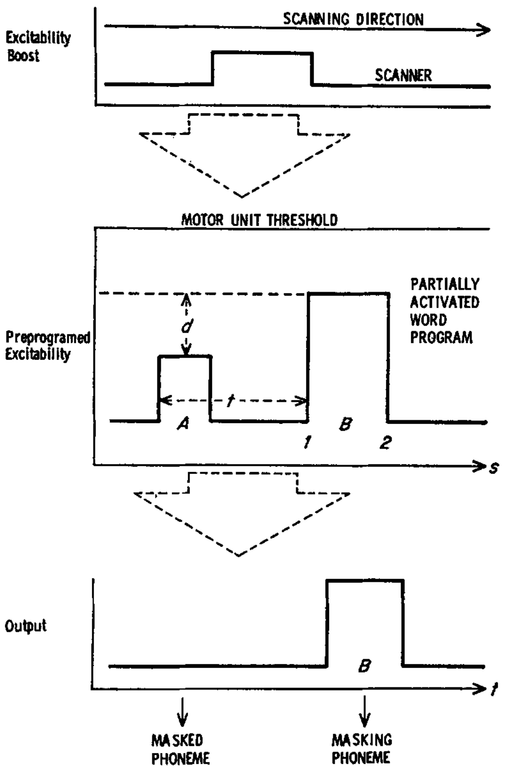

Fig. 4. An oversimplified model of stress in backward masking. The components assumed in the model are the same as those in Fig. 3 with the addition of a scanning device that sweeps over the engrams in the buffer system. The basic assumption of the model is that stressed units have a higher level of initial excitability in the buffer than unstressed units. Thus unit A in the partially activated word program in the buffer is unstressed, and unit B is stressed. Parameters D and t are discussed in the text. The broken arrows represent the process of excitability transfer in the model

(2) Stress

Our model of masking assumed that stressed units are closer to threshold than unstressed ones in the partially activated motor programme for words. Indirect support for this assumption comes from several sources. One is the fact that unstressed elements are more frequently omitted than stressed elements (Heffner, 1963). More direct evidence is found in physiological studies of muscular activity in speech production. For example, Ladefoged (1969) found increased subglottal activity accompanying stressed sounds; Klatt, Stevens, and Mead (1969) found a higher rate of airflow for stressed than unstressed segments (presumably due to increased subglottal pressure). Finally Fromkin (1968) reported an intensification of inhibitory and excitatory activity in articulators producing stressed sounds. All of these findings are consistent with the assumption that stressed units have a higher degree of subthreshold activation than unstressed units in a linear scanning system.

Returning to stuttering, if the excitability of a unit is very close to threshold and oscillates in the manner described above, the peak in oscillatory excitability could easily cross threshold, thereby repeating a just produced motor act. But the same oscillation in excitability of a unit farther from threshold might fail to reach the critical threshold value so that the motor output would not be repeated. Thus, our stress assumption predicts a lower probability of stuttering on unstressed than stressed units, which have a programmed excitability nearer threshold.

Specifically, stuttering occurs in this model when

17. $Ox > 1 - x$

where Ox is the peak increase in oscillatory excitability of a unit with preprogrammed excitability x. Thus, stuttering will be most likely when x is large i.e. when a unit is stressed, which explains the preponderance of stuttering on stressed units.

Problems for Further Research

(1) Relations between Pathological and Context-Dependent Stuttering

What differentiates the stutterer from the nonstutterer? Two possibilities are suggested in our formula for stuttering (17). One is that Ox is a physiological parameter that varies from individual to individual; the larger Ox for an individual the higher the probability he will stutter. Another possible distinguishing factor is $1 - x$. That is, the preprogrammed level of excitability of the speech motor units may be closer to threshold for some individuals, increasing their probability of stuttering.

Of course, pathological stuttering is undoubtedly more complex than stuttering in nonstutterers, since pathological stuttering may itself become a source of anxiety and perhaps even influence the stutterer's choice of words. But the parallels between stuttering in normal individuals and stutterers should not be overlooked. For example, stutterers also tend to stutter on stressed phonemes, and on syllabic initial and word initial phonemes. One also wonders whether recurring phonemes present a problem for pathological stutterers in the same way as for normal individuals.

(2) Blocking

The reciprocal inhibition assumption also suggests a possible mechanism of blocking — the transitory

5 Of course, a weaker degree of reciprocal inhibition might also obtain between slightly less similar units in a preprimed motor sequence. For example, one cannot help but wonder whether the S's underlined in 15 and 16 did not contribute to the stuttering on Sch, an identical phoneme in all but one distinctive feature.

15. Seine Sch-Sch-Schuller sind

16. s'ist sch-sch-schg'schpert

(dialectic s'ist and phonetic spelling after Meringer).

inability of stutterers to produce certain speech sounds. There has been considerable debate in the literature on stuttering as to whether blocking is "more basic" than prolongations and stuttered repetition of speech sounds, but this issue lacks substance. The important question is what mechanisms produce blocking and whether the same mechanisms would produce stuttering under certain circumstances.

There is some suggestion that the repeated phoneme effect in stuttering also holds for blocking in normal individuals. Consider example 18 from Meringer (1908):

18. Ich wollte sagen: „Wo wir wenigstens ruhig reden können," kam aber nur bis „weingstens", wo ich pausieren mußte, weil das r̲ mir nicht zur Verfügung stand.

This repeated phoneme effect suggests that blocking closely resembles masking, and results from reciprocal inhibition between the motor units coding highly similar allophones. However, one wonders why the train of speech stops in blocking but continues in masking with the omission of the repeated phoneme. One rather obvious possibility is that individuals who block closely monitor their speech and discontinue their output when the inhibited allophone fails to come forth, but further research on this question is needed.

(3) Prolongation

Prolongation refers to the involuntary lengthening of a speech sound. Consider the prolongation in 19 (speech of a normal individual, from Meringer).

19. Man sagt „die D-ido" mit Längung des Dentals. (Lengthening represented by the dash following *D*.)

Such examples suggest that repeated phonemes also play a role in prolongations. Two possible explanations of such a repeated phoneme effect are discussed below.

Feature Inhibition

Pathological stutterers sometimes stutter on a single feature of a phoneme. For example, the lip movements for *p* may be silently stuttered with neither airflow nor phonation, as if the motor units for phonation were inhibited, while those for lip movement were in a state of oscillation. This feature-inhibition hypothesis may also explain prolongations of speech sounds. Specifically, if position of articulation alone were undergoing inhibition, the articulatory configuration would remain constant while phonation and airflow continued unhampered. Thus the end result would be prolongation of the phoneme.

Fusion Model of Prolongation

Another possible explanation of prolongation can be described as a fusion model. Fusion in sensory systems is a well-known phenomenon. If a light flickers slowly, the relevant neural output of the visual system will oscillate. But if flicker is rapid enough, the system no longer follows the reverberating input; rather its output becomes continuous or fuses. It would be reasonable to expect similar fusion with rapid oscillation in excitability of motor units. Thus, when the excitability of reciprocally inhibiting units oscillates rapidly enough, output fusion might occur. Of course in speech motor systems this high level fusion would cause lengthening of a speech sound, explaining a repeated phoneme effect in prolongation.

(4) Speech Rate

The effect of speech rate presents a major problem for a general model of stuttering. Stuttering is most likely when an individual attempts to speak rapidly. Two models of this effect are discussed below:

Scanning Model of Speech Rate

The basic principle of a scanning model is as follows: output is preprogrammed in a buffer store that primes or partially activates a corresponding set of motor units (without regard for serial order). But the final speech output is determined by a mechanism independent of both the engram in the buffer and the motor units themselves. This mechanism is a scanning device that sweeps over the buffer system at a voluntarily controlled rate and in undirectional fashion, boosting the excitability of units in the buffer system to threshold in proper serial order.

Two important questions in a scanning model are the shape of the scanner and the number of units in the buffer system simultaneously covered by the scanner. In Fig. 5 we arbitrarily represented the scanner as a step function. However, our data on stuttering suggests that the function may be skewed. The fact that Proactive Induction was more frequent than Retroactive Induction is understandable if the scanner has a sharp "left" edge and a trailing "right" edge with left to right scanning.[6] In addition, if the scanning mechanism simultaneously covers at least two syllables the interactions between phonemes in adjacent syllables would be explained.

Finally, the effect of speech rate on stuttering is understandable in a scanning model that represents stuttering as the result of oscillation in excitability around the threshold of a low-level motor unit. If the rate of scanning is sufficiently slow, supra-threshold excitability of this low-level unit may be maintained from above for a period longer than its period of oscillation, so that the oscillation would not become manifest in articulation. And since a slow rate of scanning implies a slow rate of speech, this model would explain the reduced probability of stuttering at slower rates of speech.

Gamma Loop Model of Speech Rate

Another explanation of the effect of speech rate on stuttering is possible should speech rate turn out to be determined at a low level as by the time settings for a gamma loop system. In this case the articulatory configuration could be held constant by the low level

6 Of course, the spatial notions "left to right" are only a manner of conceptualizing the problem, and in fact are irrelevant to both the mathematical description and physiological aspects of a scanning mechanism (see MacKay, 1969, for elaboration of this point).

gamma system during the period of oscillation of the high level (alpha) component that originally instigated the articulatory configuration. And since the gamma system would hold an articulatory configuration longer at slower rates of speech, this model would also explain the effect of speech rate on stuttering.

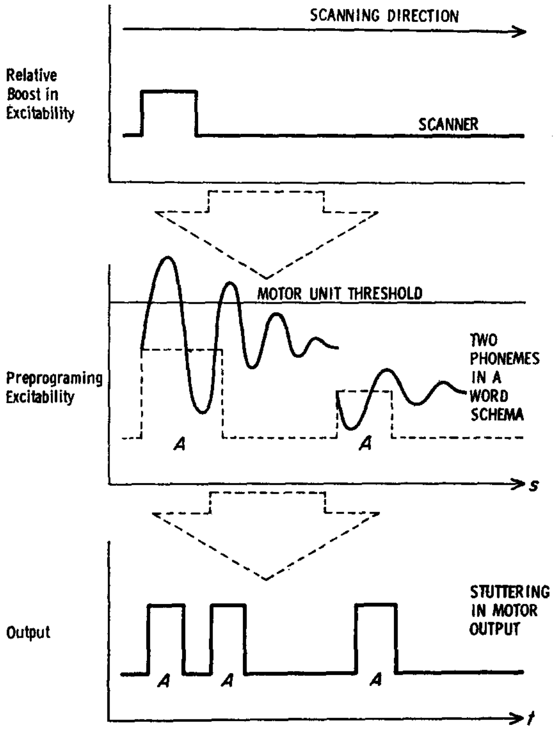

Fig. 5. A model of stress in retroactive stuttering. The unit on the left in the word schema is stressed and the other is unstressed; these units are represented as highly similar and so are assumed to interact in reciprocal inhibitory fashion at the motor unit level (see Fig. 3). Thus excitation of one component (in scanning) will bring about oscillation in excitability of both units. Oscillation in the buffer program itself (see Fig. 3) is not assumed, but rather in the motor units for producing the contextual variants of phonemes. But since the excitability from priming (i.e. display in the buffer), scanning, and oscillation are assumed to add, the end result can be represented at a single level. Thus oscillation is shown by the solid line superimposed on the preprogrammed excitability. When the peak in oscillatory excitability exceeds the threshold of the motor units, stuttering in the output will occur according to this model. Since stressed units are closer to threshold, than unstressed units, peaks in excitatory oscillation will exceed threshold more often for stressed than unstressed units, explaining the preponderance of stuttering on stressed phonemes

(4) Synonymic Intrusions and Stuttering

Hocket (1965) postulated an interesting relation between certain instances of stuttering and synonymic intrusions. Synonymic intrusions are defined as involuntary mixtures or blends of synonymic words. For example, 20 is a meaningless combination of 21 and 22. Here the individual began to say auffallend but switched in midstream to anstößig (from Meringer, 1908).

20. aufstössig (non-word)
21. auffallend (shocking)
22. anstössig (shocking).

Similarly, 23 (nonsense word) was an involuntary combination of 24 and 25, synonyms for evaporated (from Meringer, 1908).

23. verdumpft (non-word)
24. verdampft (evaporated)
25. verdunstet (evaporated).

Hocket's theory was that fear of producing such synonymic fusions may result in stuttering. Specifically, "a word can be broken off virtually as it begins for fear that it will not come out right, and then indecision between one word and another or between the right word and a threatening blend can produce a series of attempted corrections, each in turn cut off" (Hocket, 1965). Thus, fear of synonymic intrusion in Hocket's model may cause an individual to stop and start again — the end result being stuttering.

Meringer documents several cases of intrusion stuttering (e.g. 26).

26. „Ein R-R-Roastbeef" sagte er weil Rindfleisch mit in Gedanken neben Roastbeef vorhanden war.

Here one wonders whether the stuttering on R in Roastbeef could have been induced by the R in the same lexical and syllabic position in the interfering *Rindfleisch*. Similarly, in 27 the W in the interfering *Weib* may have induced the stuttering on the W in *verwaist*.

27. „Sie sind heute verw-w-waist" sagte er. Ich begriff den Grund des gestotterten W̲ und frage ihn woran er nebenbei gedacht hat. Er sage „An Weib!" Er meinte, „ohne Weib".

This example again suggests that repeated phonemes may play a role in stuttering induced by threatened blends.

Viewed from the framework of the present model, blends *per se* suggest that the semantic component in speech production tends to be somewhat nonspecific, calling up several words having roughly the same meaning. And although one of these programs is finally selected for scanning, the subordinate programs in the buffer system still prime the motor units. Thus, when similar allophones are simultaneously primed, as in 27 where *verwaist* and *Weib* simultaneously occupy the buffer store, stuttering would occur in this model in the same way as context-dependent stuttering. Thus, our model explains both intrusion stuttering and context-dependent stuttering with the same principles. But further research is needed to test whether this or Hocket's model best describes intrusion stuttering.

In conclusion we have outlined a causal model of stuttering at the phonetic level. By a causal model we mean one that accounts for an event in terms of the factors that might cause it, exclusing "catalytic factors" indirectly hampering or facilitating the action of a causal factor. For example the grammatical or semantic function of words is known to influence the probability of stuttering (Soderberg, 1967). However we must be able to envision how such factors could directly bring about repetitions, blocking or prolongations before including them in a causal model. Or taking another example, a rapid speech rate may increase the probability of stuttering (in either of the ways described above) without in any way being a necessary or sufficient condition for stuttering. This distinction between causal and catalytic factors must not be confused with that between direct and remote causes. That is, if A causes B and B causes C then in reference to C, B is a direct cause and A is a remote cause. But both A and B are causal and not catalytic factors. Of course, a general model must eventually specify how any known catalytic factor influences a proposed causal mechanism. The ultimate strength of our model will lie in its ability to assimilate catalytic factors known to influence stuttering.

Finally, the limitations of the data we have considered in constructing our model should be pointed out. On the one hand our analyses lacked the control over variables such as speech rate and syntax that could be obtained in experimental tests of our model. But on the other hand, data on spontaneous speech are sure to reflect the normal mechanisms of speech production and thereby lay the foundation for further experimental studies.

Summary. An analysis of stuttering in the natural speech of Germans revealed the following facts:

1. Stuttered phonemes are frequently preceded or followed in the context by an identical phoneme defined as the inducing phoneme.

2. The inducing phoneme usually followed rather than preceded the stuttered phoneme.

3. The inducing phoneme occurred closer to the stuttered phoneme than would be expected by chance.

4. The stuttered and inducing phonemes were usually situated in identical syllabic positions.

5. Stuttered phonemes usually occurred in stressed syllables.

Three main assumptions seemed necessary to explain these findings:

1. Both syllables and phonemes are units in a hierarchy of speech motor determinants.

2. Contradictory aspects of similar motor programs (at both the syllabic and phonemic levels) interact in reciprocal inhibitory fashion. This assumption also provided a possible explanation of blocking and prolongation of speech sounds — phenomena which occur in contexts similar to stuttering.

3. The subthreshold excitability for stressed units is greater than for unstressed ones.

Independent support for each of these assumptions was advanced. Thus, the contextual structure of stuttered speech provided a possible solution to the riddle of stuttering.

References

Brown, J. L.: Vision and visual perception, ed. Graham, C. H. New York: Wiley 1965.

Brown, S. F.: Stuttering with relation to word accent and word position. J. abnorm. soc. Psychol. **33**, 112—120 (1938).

Bullock, T. H.: Mechanisms of integration, chapter 5 in Bullock and Horidge. Structure and function in the nervous systems of invertebrates. London: Freeman 1965.

Eisenson, J. (ed.): Stuttering: a symposium. New York: Harper & Row 1958.

Hartline, H. K., Ratliff, F.: Inhibitory interaction of receptor units in the eye of limulus. J. gen. Physiol. **40**, 357—376 (1956).

Heffner, R. M. S.: General phonetics. Madison: University of Wisconsin Press 1964.

Holst, E. von, St. Paul, U. von: On the functional organization of drives. Anim. Behav. **11**, 1—20 (1963).

Hubel, D. H., Wiesel, T. N.: Receptive field and functional architecture in two nonstriate visual areas of the cat. J. Neurophysiol. **28**, 229—289 (1965).

Klatt, D. H., Stevens, K. N., Mead, J.: Studies of articulatory activity and airflow during speech. In: Bouhys, E. (ed.), Sound production in man. Ann. N.Y. Acad. Sci. **155**, 1 (1968).

Lashley, K. S.: The problem of serial order in behaviour. In: Jeffress (ed.), Cerebral mechanisms in behaviour, p. 112—136. New York: John Wiley & Sons 1951.

Levinson, J. Z.: Flicker-fusion phenomena. Science **160**, 21—28 (1968).

MacKay, D. G.: Spoonerisms: The anatomy of errors in the serial order of speech. Neuropsychologia (1969) (in press).

— Forward and backward masking in motor systems. Kybernetik **2**, 57—64 (1969).

Meringer, R.: Aus dem Leben der Sprache. Berlin: Behr's Verlag 1908.

— Mayer, K.: Versprechen und Verlesen: Eine Psychologisch-Linguistische Studie. Stuttgart: Göschen 1955.

Ratliff, F.: Inhibitory interaction and the detection and enhancement of contours. In: Rosenblith, W. A. (ed.), p. 183—203. Sensory communication. New York: Wiley 1961.

Sperling, G.: A model for visual memory tasks. Hum. Factors **5**, 19—31 (1963).

Stetson, E. H.: Motor phonetics: A study of speech movements in action. Amsterdam: North Holland Publ. 1951.

Teuber, H. L.: Perception, chap. 65, in: Field et al. (eds.), Handbook of physiology III. Neurophysiology. Amer. Physiol. Soc., 1595—1668 (1960).

Wickelgren, W.: Distinctive features and errors in short-term memory for English consonants J. acoust. Soc. Amer. **39**, 2, 388—398 (1966).

Wyneken, C.: Über das Stottern und dessen Heilung. Z. rat. Med. **31**, 1—20 (1968).

D. G. MacKay
Speech and Communication Research Unit
University of Edinburgh
31 Buccleuch Place
Edinburgh, England