INVESTIGATION OF SPEECH PRODUCTION PROCESSES

By

ROGER WALTER HARNISH

Bachelor of Arts

University of Rochester

Rochester, New York

1970

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
MASTER OF SCIENCE
December, 1973

INVESTIGATION OF SPEECH PRODUCTION PROCESSES

Thesis Approved:

Robert J Weber
Thesis Adviser

Robert F Stanners

Larry Hochhaus

Cheryl M. Scott

N N Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

Look to this day!
For it is the very life of life.
In its brief course
Lie all the verities and realities of your existence:
The bliss of growth
The glory of action
The splendor of achievement,
For yesterday is but a dream
And tomorrow is only a vision,
But today well lived makes every yesterday a dream of happiness
And every tomorrow a vision of hope.
Look well, therefore, to this day!
Such is the salutation to the dawn. (Anonymous)

I would like to thank the members of my committee, Dr. Robert Stanners, Dr. Larry Hochhaus, and Dr. Cheryl Scott for their interest, helpfulness, and concern.

I would like to thank my committee chairman, Dr. Robert Weber, not only for the tremendous amount of help he has given me but also for being a special kind of person.

I would like to thank two good friends for their help, Marilyn Mallue and Mike Wegmann.

And a special thanks and all my love to Marty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

LITERATURE REVIEW

Introduction

The intent of the present paper is to investigate the role of motor processes in the control of speech production. Given a simple unit of analysis, attention can be focused and complicated phenomena become describable as lawful compounds. Therefore, the present study attempts to focus on "speech motor activity" as a unit of analysis. It is realized, of course, that speech production and control can involve a tremendous number of such units and a tremendous number of complex interactions between these various units. However such a comprehensive evaluation is beyond the scope of the present paper. Thus, it is hoped that the boundaries describing the "speech motor activity" unit will be clear enough to isolate the present analysis at this level.

Briefly, the present evaluation of the role of motor control was performed as follows. Experimental sets (instructions) were used to manipulate speaking conditions so that variations in speech motor activity could be observed. Predictions as to expected motor responses were generated and evaluated against the experimental results. Predictions were based on two general theoretical perspectives: (1) Grammatical-Lexical based models, and (2) Spatial-Target based models.

A problem concerning the attempt to isolate only those aspects of of the speech process involved with motor activity should be noted. Response time is the dependent variable for the experimental manipulations. If intrusions from other control units impinge on motor functioning, and if this effect is not constant and uniform, then distortion of response time can occur. Any distortion in response time would most likely make clear-cut evaluation of the speech motor hypotheses impossible. An example of impingement on motor functioning is given by Weiss (1964). Weiss observed that the rate of speech of a person's repetition of sounds, as in stuttering, can be "the exact syllabic speed of his nonrepetitive (free-flowing) speech....A faster or slower rate indicates that the individual has become aware of his repeating and is attempting to correct it" (p. 20). Thus, awareness appears to be indicative of some unit of speech processing impinging on motor unit control.

To control possible intrusions, instructional sets were instituted. In a pilot study, subjects were instructed not to correct speech errors and not to pause while speaking. Then the subjects were given difficult phrases (tongue twisters) to repeat. The results indicated that intrusions were eliminated in approximately half the subjects. In the present study stronger and more precise admonitions against intrusions were given in an attempt to achieve total elimination of the intrusions. As a check on the success of the instructional sets in eliminating intrusions, an "intrusion checklist" was devised which catalogued evidence of (a) speech error correcting, (b) rate or response changes, and (c) any other miscellaneous forms of intrusion.

History and Description of Speculations as to the Nature of Motor Control in Speech Production

The first major theoretical concern about the nature of central and peripheral speech control processes was shown by Lashley (1951). Lashley was concerned with how people produce smooth, temporally ordered sequences of sounds. He concluded that "elements of a sentence are readied or partially activated before the order is imposed upon them in expression" (Lashley, 1951, p. 535). This idea is still considered accurate today. As to the "selective" mechanism by which the various sounds, words, or utterances are picked out and ordered, Lashley admitted that he had no answer.

It has been in search of this answer that subsequent research has continued where Lashley stopped. The earliest theories and models of speech production stemming from Lashley's work were based primarily on observable grammatical-lexical performances. (Hereafter, this will be referred to as the Grammatical-Lexical perspective.) These observations resulted in theories that rested strongly on the idea of structure and organization--that is, rule-governed behavior. As this perspective developed, it became necessary to define hypotheses about what the rules might be governing. Thus the idea of invariant linguistic units (such as the phoneme, the syllable, and the word) became important.

For example, the concept of the phoneme is very useful in the study of language as behavior. As MacNeilage (1970, p. 182) points out:

> The content of every known language can be described (with good, though not perfect interobserver agreement) as being made up of a finite and relatively small number of phonemes. Furthermore, the main features of the sound pattern of a given language can be economically described by laws of combinations of its phonemes. In addition, the construction of the alphabet is predicated on the assumption that a language consists of combinations of a limited number of distinctive sound units of phoneme size.

However, all this information is mostly observational, not experimental. Therefore, it was felt that if the experimental behavioral reality of invariant units could be established, then the assumption of speech control being a function of "rule-governed permutations of a finite set of invariant unit commands" (MacNeilage, 1970) would definitely be supported (this is not to say that other theories could not also derive support from these findings). The above assumption also presupposes that any lack of invariance between a speech motor response and a linguistic command is due to "noise" in the system.

There is direct evidence for the behavioral reality of the phoneme (Liberman, 1957; and Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). There is also evidence for the behavioral reality of the syllable, the word, and other linguistic features (MacKay, 1972). Supported by this evidence, many Grammatical-Lexical models of speech production such as Liberman et al. (1967); Hockett (1967); Fromkin (1971); and MacKay (1972) have recently been developed. Some of these models are based on behavioral data collected by observing speech errors (generally defined as "unintentional linguistic innovations" or "involuntary deviations in performance from the speaker's current phonological, grammatical, or lexical intention" [Fromkin, 1971, pp. 28-29]). Other models are based primarily on examinations of the physiological possibilities and constraints of the vocal apparatus. Both views turn out to be relatively similar and both find support within their respective fields (Fromkin, 1971; Liberman et al., 1967).

There are many differing opinions as to the <u>specific</u> speech control functions of speech motor activity; however, there is fairly good agreement as to the <u>general</u> control functions. The general agreement is that speech processing proceeds (a) in accordance with abstract hierarchiacal rule structures, (b) by means of parallel articulo-motor operations on subphonemic features of a given phoneme, and (c) in relation to the prior as well as the present configuration of the articulatory system (Liberman et al., 1967; MacKay, 1972). The parallel subphonemic feature processing results in efficient speech production that allows articulators to perform at a reasonably high speed. The parallel processing interacting with the merged effects of past and present instructions on the configuration of the articulatory system results in temporal and spatial overlaps of the component parts. These overlaps produce a very complex relationship between phonemic specifications and articulatory movements. In spite of this complexity, the relationship is felt to be predictable and invariant for the most part. However, some variability between phonemic specifications and articulatory movements in the form of inefficiency or noise would not be out of the question.

MacNeilage (1970) discusses experimental investigations which have attempted to support the foregoing conclusions by demonstrating that the electromyographic (EMG) correlates of the phoneme are invariant to a significant extent. He concludes that the studies support this concept of invariance to some extent, although more often than not "ubiquity of variability" rather than invariance has been demonstrated.

A major criticism of the Grammatical-Lexical perspective has been offered by MacNeilage (1970). The criticism is in the form of an argument, namely that the basic question in speech production theory should be, "How do articulators come as close to reaching the same position as they do" (p. 184)? MacNeilage points out that theorists who accept the priority of the assumption of discrete invariant units (Grammatical-Lexical theorists) necessarily ask the incorrect question--why do articulators not always reach the same position for a given phoneme (p. 184)? Thus, they attribute

> apparently contradictory peripheral variability to inefficiency or 'noise' in the executive mechanism, rather than to the operation of important control mechanisms (p. 184).

The above criticism of the Grammatical-Lexical perspective has become the point of departure for a new trend in studying speech production control, namely the Spatial-Target perspective. Target theories attempt to explain what Grammatical-Lexical theories attempt to ignore, namely how variability of muscular responses can occur and yet still produce invariant vocal responses. The central belief is that the speech production process is not an inefficient response to invariant central signals, but is "an elegantly controlled variability of response to the demand for a relatively constant end" (MacNeilage, 1972). Basic to this belief are the concepts of motor equivalence, space-coordinate systems, and spatial targets. Speech production originates with the reception by a space-coordinate system of information concerning the utterance required. This system, which controls the vocal motor tract, then converts the information into one or more spatial target specifications. This process results in demands on the motor system control mechanism to generate movement command patterns which, in turn, allow the articulators to reach the specified target or targets.

Speculations as to the overall mechanics of spatial target control have centered on neuro-electrophysiological explanations (MacNeilage, 1970; Sussman, 1972). Specifically, investigators have hypothesized that neuro-loops control the speech motor production process. Of concern in the present paper are four neuro-electrophysiological hypotheses: (1) an open loop controls the targets, (2) a closed loop controls the targets, (3) a dual open loop-closed loop operation controls the targets, (4) a dual closed loop-closed loop operation controls the targets. Before discussing these four hypotheses, some definitions are presented.

An open loop system is defined as follows (Mysak, 1966, p. 6):

> Open loop systems describe devices which carry out a series of operations in a certain prescribed manner and which do not possess the potential for changing their operations in instances where the results are not those desired. Time-operated traffic lights and timing mechanisms in various home appliances serve as good illustrations of this type of automatic control.

A closed loop system is defined as follows (Mysak, 1966, p. 6):

> Closed loop systems are different from open loop systems inasmuch as they are error sensitive, error measuring, self-adjusting, goal-directed mechanisms. These systems operate to control the machine of which they are a part. They feed back into the machine information pertaining to its performance and, thereby, effect automatic corrections whenever error-performance signals are received. Familiar mechanisms of this type include home heating, water heating, and refrigerator thermostats.

An open loop-closed loop system is defined simply as an open loop system operating in conjunction with a closed loop system to control two different aspects of the same operation. An illustration would be the combination of a time-operated traffic light and pressure operated road switches that recycle the green light to the desired traffic lane.

A closed loop-closed loop system is defined as one closed loop system operating in conjunction with a second closed loop system to control two different aspects of the same operation. An illustration would be a parking lot gate. Upon entering the lot, a coin in the machine will raise the gate, and after driving through the entrance a pressure device on the lot will indicate the car is through the entrance and will lower the gate.

The Open Loop Hypothesis

The open loop hypothesis indicates that the motor system control mechanism might not have to wait for information to actually reach a previously specified location before being able to control the next movement appropriately. In other words, it is likely that the open loop control mechanism preprograms movement command patterns a number of units (phonemes) ahead of their utilization as muscle contractions. One way of characterizing open loop control is to state that it is particularly well equipped to handle rapid speech production.

No direct behavioral or neurological evidence of open loop control of speech has as yet been demonstrated. However, there is some suggestive data indicating the possibility of open loop control. An unpublished experiment by MacNeilage, Krones, and Hanson (in MacNeilage, 1970) has demonstrated a significant positive correlation between the velocity of opening the jaw for a sound and the velocity of its closing after termination of the sound. The experiment has also demonstrated a significant positive correlation between the velocity of jaw openings and the maximum amount of a jaw opening. MacNeilage (1970) feels that these findings can be interpreted in terms of "a single open loop dynamic control process" (p. 192). (It should be noted that Sussman /1972/ uses this same experiment to support the existence of closed loop control.) Weiss (1964) has found that in normal speech, the speaker inhales the exact necessary amount of breath so that he neither runs out of breath nor is left with an excess of breath when pronouncing the sentence length he intends to speak. Weiss has also found that the first notes of a person's speech melody indicate the length of the coming sentence or phrase. Both of these phenomena are examples of preprogrammed (as opposed to self-adjusting) operation and thus may be interpreted in terms of open loop control. Finally, Fromkin (1971) notes that studies of errors in speech show that the distance between the initiation of a possible syllabic-substitution error and the actual substitution does not generally exceed seven syllables (e.g., phi-so-lo-phy for phi-lo-so-phy). Since the short-term memory span of man is approximately seven units (Miller, 1956), one might assume that if open loop control exists preprogramming would run about seven units (phonemes or syllables) per generated command pattern. Thus, a speech production system in which units of speech were always displaced within seven units from their initiation would be consistent with an open loop control explanation.

The question now is why should open loop control be the best and/or only means of controlling speech production. MacNeilage (1970,

p. 190) feels the answer is that "the speech motor mechanism appears ...to be very favorably equipped for such a mode of operation". To make this point clear, MacNeilage compares speech to tennis. He notes that a tennis player is able to anticipate, to some extent, the position of his target (the ball). However, "In speech, anticipation of targets can extend much further into the future than in tennis" (p. 190). Also, he notes that whereas in tennis the targets are extrinsic and dependent on an extrinsic response (the immediately previous response of the opponent), in speech the targets are intrinsic and "have a fixed spatial relation to the structures to which the muscles, themselves, are mechanically coupled" (p. 190).

In criticism of the open loop hypothesis, MacNeilage (1970) notes that any unit (a phoneme) can be preceded by any number of other units (any 1 of 20 other phonemes) each making unique mechanical demands on any articulator involved in the unit-to-unit transition. To account for all possible phonological combinations (remembering context and stress patterns are involved) storage and availability of over 100,000 control operations would be needed. In MacNeilage's opinion, "it becomes obvious that open loop control of this system [because it must preprogram the control operations] would require the storage and availability to the motor system of an enormous amount of *a priori* information" (p. 190).

The Closed Loop Hypothesis

The closed loop hypothesis indicates that closed loop (feedback) control circuits could constantly sample the mechanical state of an articulator and adjust motor commands to fit previous locations and states of the articulator to target specifications. This type of closed loop is a "difference" feedback loop. The output is matched to the input, and speech production is not continued until the difference between the output and the input is minimized to the proper criteria. One way of characterizing closed loop control is to state that it is particularly well equipped to handle accurate speech production.

There is some fairly direct behavioral and neurological evidence of the existence of closed loops (Sussman, 1972; MacNeilage, 1970). Sussman (1972) has found closed loop feedback mechanisms within the tongue and other parts of the speech musculature that are capable of keeping higher order neural centers constantly aware of the spatial position, the direction of movement, and the rate of movement of the articulators. In addition, Sussman (1972, p. 264) reports a study in which the gamma-efferent nerve fibers innervating the muscle spindles of the jaw were selectively anesthetized. The movement records showed gross loss of fine positional control, a consistent reduction in jaw velocity and acceleration, and a significantly altered pattern of jaw movements. These findings were taken as stressing the importance of normally operative muscle-spindle receptors if finely coordinated and precise neuromotor control of the speech musculature is to take place. MacNeilage (1970, p. 191) reports two studies which give evidence suggesting that gamma motor loops can be brought under voluntary control by human subjects. Also, Weiss (1964) reports the indigenous savage populations of Africa and Australia, if asked to report on it, can indicate fairly precisely the position of their tongue, mouth, pharynx, and so forth during ongoing articulation.

Why should closed loop control be the best and/or only means of controlling speech production? MacNeilage (1970, p. 192) feels that the answer is:

> because of the enormous amount of a *priori* information the brain appears to require for open loop control, and because of the considerable indirect neurophysiological evidence that possibilities for closed loop control exist.

In criticism of the closed loop hypothesis, MacNeilage (1970, p. 192) states that "although there appears to be closed loop control of some aspects of initiation of speech, it is not known at present whether closed loop control operates within utterances". He also points out that studies using anesthetics, while having shown deterioration in the quality of speech, have not ruled out the possibility that some, if not all, of the deterioration might be due to gross changes in the functional balances of neural activity.

The Open Loop-Closed Loop Hypothesis

The open loop-closed loop hypothesis indicates that the open loop component emits preprogrammed context-independent commands for articulators to reach certain positions, and that the closed loop component constantly monitors and adjusts the commands as necessary. There is no direct behavioral or neurological evidence for the existence of an open loop-closed loop control mechanism. However, since arguments and support for both open loop control and closed loop control have been rendered (although not decisively in either's favor), this could mean that the non-decisiveness is due to the fact that the two loops are both present and working together in some fashion. Given the possibility of open loop-closed loop control, one way of characterizing this control would be to state that it would be particularly well equipped to compromise between rapid and accurate speech production.

The Closed Loop-Closed Loop Hypothesis

The closed loop-closed loop hypothesis indicates that a dual system of closed loop feedback control over speech production is operating. One loop could control initial gross articulator placement and coordination (possibly by means of the alpha system). The second loop could take over from the first loop and control the more precise placement and coordination of the articulators (possibly by means of the gamma system).

The reason for speculating on dual controls for closed loop feedback is that it appears more consistent with the neurological evidence known today. For example, MacNeilage (1970, p. 191) in referring to negative evidence of single closed loop control, points out that gamma control of running speech requires a different set of commands being issued about every 70 millisecs. He notes, "such rapidly varying phasic control has not in the past been associated with the gamma system". Controlling initial gross movement commands by the alpha system might take some of the time pressure off the gamma system and thereby allow it to control the precise movements of speech as Sussman (1971) indicates it does. Also, as was noted earlier, studies using anesthetics demonstrated that precise motor control of speech was impaired while gross motor control remained operable (Sussman, 1972). Again, this could imply two closed loop feedback systems--one gross, one precise. Finally, one of the characteristics of stuttering or cluttering speech behavior is the idling phenomenon (Weiss, 1964). An utterance is initiated and then "idles" or repeats itself as in stuttering. Weiss (1964, p. 19) says that this idling phenomenon is the same as speaking without having access to words. This phenomenon could easily be explained by the closed loop-closed loop hypothesis as follows. The gross closed loop control is operative causing the idling while more precise closed loop control is temporally unable to take over, thus causing the lack of "word fuel". Alternatively, this idling phenomena would seem to be very difficult to explain by the single closed loop control hypothesis or the single open loop control hypothesis.

In summary, the foregoing hypotheses were intended to speculate only on speech motor control activity. It is readily acknowledged that given a broad enough scope, any aspect of the speech production system can be described as either an open or closed loop or both. Hopefully by sticking closely to a motor control unit level of analysis, the distinctions between the various loops can remain meaningful. Along the same lines, it is possible to attribute open and/or closed loop functioning to the Grammatical-Lexical model; it indeed displays characteristics of both concepts and any differentiation may only be in the eye of the beholder. However, the Grammatical-Lexical model is unique in that it proposes an invariant relationship between input to the speech motor production process and the output in the form of motor movement and, in addition, proposes no flexibility for the execution of this relationship. On the other hand, the target models propose that input and output have an invariant relationship and that there is flexibility in the means of producing this relationship (either by variable methods of preprogramming movements and/or by in-process modifications).

A summary of the five speech control hypotheses is presented in Table I.

TABLE I

A SUMMARY OF THE SPEECH CONTROL HYPOTHESES

Hypothesis	Control Characteristics	Possible Function
Grammatical-Lexical	Invariant relationship between phonemic specifications and articulatory movements. Parallel processing of sub-phonemic features.	Rapid and accurate speech
Open Loop	Preprogramming of several speech units at a time.	Rapid speech
Closed Loop	In-process feedback. Input and output must reach a critical small difference before speech can continue.	Accurate speech
Open Loop-Closed Lood	Preprogramming of several speech units at a time. In-process feedback. Input and output must reach a critical small difference before speech can continue.	Compromise as needed between fast and accurate speech
Closed Loop-Closed Loop	In-process feedback. Difference criteria for first loop fairly sizeable. Difference criteria for second loop fairly small.	Smooth accurate speech

CHAPTER II

STATEMENT OF THE PROBLEM

The Basis for Making Predictions when the Dependent Variable is Speech Response Time

The purpose of the present study is to test the validity of the five speech control hypotheses which are summarized in Table I on page 15. This testing was accomplished by comparing experimental response patterns with patterns predicted by the five hypotheses. The experimental patterns were based on the response rates of implicit and explicit speech for difficult (tongue twister) and easy (non-tongue twister) phrases, under two limiting speaking conditions ("speak fast" and "speak perfectly"). The predicted patterns were based on the following considerations.

Twister Phrases (T) Versus Non-Twister Phrases (NT)

Two possible consequences of varying the difficulty of spoken material might be expected: (a) difficulty of the spoken material does not affect speech response time, and (b) more difficult material is spoken more slowly. The latter expectation is supported by experimental findings. Novikova (1961) has found that lingual activity, as recorded by EMG, increased with increasing difficulty of the thought operation, with increasing illiteracy of the participant, and with increasing unpreparedness of the participant. Locke (1970) has noted that difficulties in phoneme articulation retard speaking rate.

Fromkin (1971) has reported that repeated phonemes make speech more difficult. Therefore, T phrases might be expected to be more difficult than NT phrases. Also, it should be noted that there is a folk-history that tongue twisters are difficult to say. During motor control of speech production, this difficulty might be expressed by some articulatory switches or selection (such as ba-bla) being more difficult (and therefore slower) than others (such as ba-ha). In summary, the experimental predictions in the present paper are based on the assumption that T phrases are more difficult than NT phrases and therefore are processed more slowly. Also, differential predictions for the five speech control hypotheses are possible only if this assumption holds.

Implicit Versus Explicit Speech Responses

The relationship between implicit speech and specific functions of the speech production process are not precisely known. Three relationships are possible. In describing these relationships, speech motor activity has been divided into two parts: (1) in-process speech motor activity (motor activity has begun but no sound is yet possible), and (2) end process motor activity (motor activity has increased) to the point where sound is produced and/or motor activity has been signalled to begin producing sound). This distinction is not arbitrary but is based on experimental findings by Sokolov (1972). Sokolov reports that by neuro-electrical recording of speech motor activity he has found that before "thoughts" become expressed as explicit speech, they are preceded by a discharge of "motor speech impulses". Also, "motor speech impulse" discharge is "always antecedent to the utterance of words, be it even a matter of fractions of a second" (Sokolov, 1972, p. 66). The three proposed relationships between implicit speech and the specific functions of the speech production process are as follows.

(A) Implicit and Explicit Speech as the End Process of Speech Motor Activity. If both implicit and explicit speech are the end process of motor activity during the speech production process, then variable speech motor control activity should produce no differential responses, and implicit and explicit speech should have identical response times. This viewpoint is depicted in Figure 1.

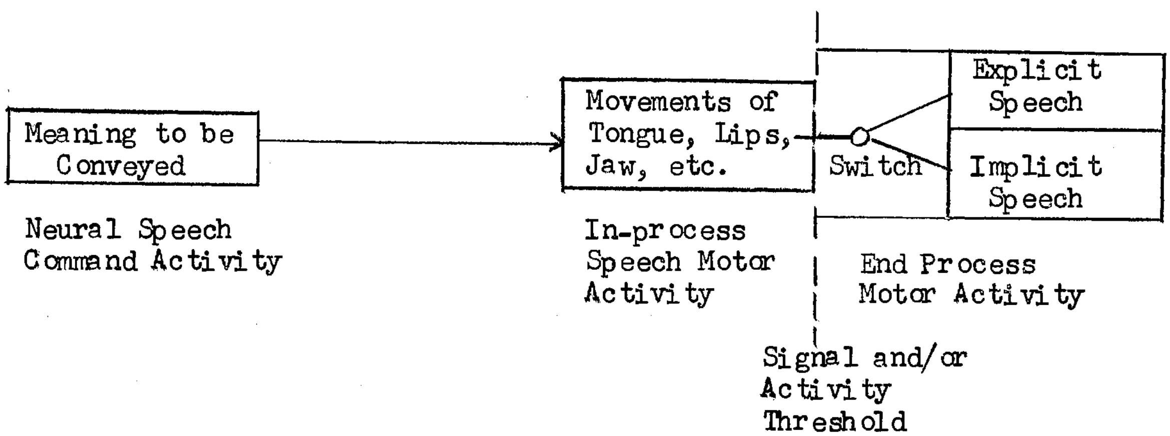

Figure 1. Flow Chart Representing Implicit and Explicit Speech as an End Process of the Speech Production Process

This view derives support from experimental work in the area of implicit and explicit speech rate production. The work by Landauer (1962), Weber and Bach (1969), and Weber and Castleman (1970) has shown that the processing times for implicit and explicit speech responses are identical. This view also derives some support from work in the area of subvocal speech, much of which has demonstrated a strong correlation between "implicit speech" and "articulatory aspects of language-related behavior"--in particular, explicit speech. (For a review of the subvocal speech literature, see Locke, 1970.) However, Locke (1970) cautions:

> A fundamental question which must be asked of subvocal speech concerns the degree to which it resembles overt articulation. The theoretical formulations in this paper assume some relationship. But previous attempts to correlate electromyographic tracings from overt and covert articulation have not been totally successful, and lacking this or similar evidence, researchers have tacitly assumed such a relationship, by implication from existing data (p. 12).

In summary, the "implicit and explicit speech as an end process" view would lead to the prediction that, given the same material and the same speaking conditions, implicit and explicit speech would always have the same response rates. Therefore, if this view is valid, no differential predictions can be made among the five hypotheses based on implicit and explicit speech responses.

(B) In-process Speech Motor Activity as Implicit Speech. If implicit speech is speech motor activity, then (assuming that explicit speech is the end process of speech motor activity) implicit speech responses will always occur before the explicit speech responses. This viewpoint is depicted in Figure 2.

This view derives much of its support from the same experimental literature that was used to support the "implicit and explicit speech as an end process" view. Subvocal speech studies all seem to indicate a relationship between articulatory activity and "implicit speech" (Locke, 1970). In addition, Landauer (1962) although finding no statistical difference between implicit and explicit speech, did note that "thinking" (implicit speech) responses take slightly less time at first than explicit speech responses. Finally, Sokolov (1972) has noted a dependent yet temporally discrete relationship between articulatory motor activity, implicit speech activity, and explicit speech activity. His experiments show that before "thoughts" become expressed as explicit speech they are preceded by a discharge of "motor speech impulses", and that both "the motor speech impulses" and the "thoughts" (implicit speech) are "always antecedent to the utterance of words, be it even a matter of fractions of a second" (Sokolov, 1972, p. 66).

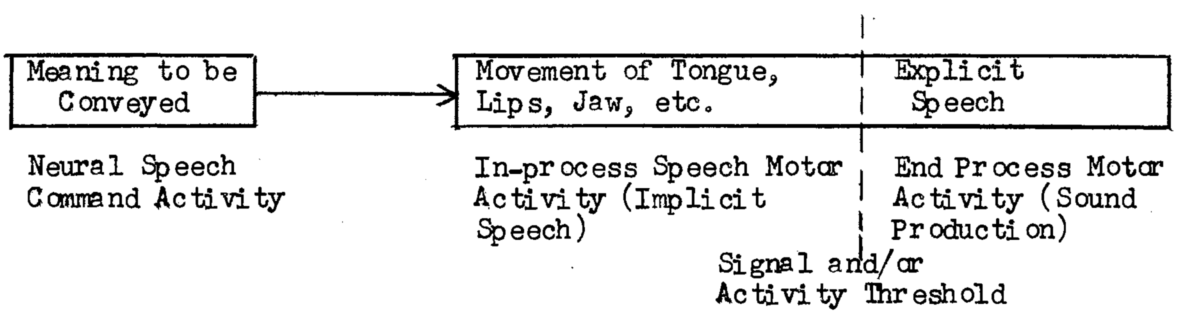

Figure 2. Flow Chart Representing Implicit Speech as In-Process Speech Motor Activity

In summary, the "in-process speech motor activity as implicit speech" view leads to the prediction that implicit speech responses will always be slightly faster than explicit speech responses.

<u>(C) Neural Speech Command Activity as Implicit Speech.</u> If implicit speech is neural speech command activity, then implicit speech responses should occur before any form of speech motor activity and consequently before explicit speech responses. This view is depicted in Figure 3.

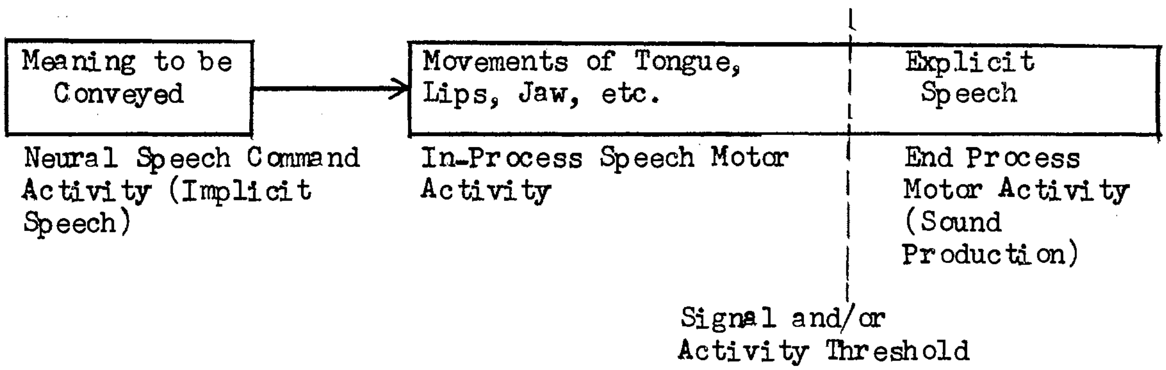

Figure 3. Flow Chart Representing Implicit Speech as Neural Speech Command Activity

This view derives theoretical support from Vygotsky (translated by Hanfmann and Vakar, 1962). Vygotsky believes that thought does not have an automatic counterpart in words and that there must be a transition from thought to word, a transition that "leads through meaning" (Hanfmann & Vakar, 1962, p. 150). This view derives experimental support from work by Dodge (reported in McGuigan, 1966, p. 60). Dodge has noted that anesthetizing the tongue and lips does not disturb implicit speech although explicit speech is markedly disturbed. Dodge feels that this implies that implicit speech is independent of (at least some) peripheral speech activity.

In summary, the "neural speech command activity as implicit speech" view leads to the prediction that implicit speech responses will always be faster than explicit speech responses, and that the difference is a function of the difficulty of the motor processing--the more difficult the processing the greater the response time difference.

Fast Speech Versus Perfect Speech

The "Fast" and "Perfect" speech sets help allow for differential prediction among the five speech control hypotheses. The Fast set should favor the speech production systems designed to handle rapid speech production while causing some need for compensation in systems designed to handle accurate speech production. The reverse should be true for the Perfect set. The Perfect set should favor the speech production systems designed to handle accurate speech production, and cause some need for compensation in systems designed to handle rapid speech production.

Intrusion Scoring

A checklist was devised to ascertain the frequency of intrusions. The checklist contains seven intrusion categories which represent objective evidence of (a) speech error correcting, (b) response changing, and (c) any other miscellaneous forms of intrusion. Each time a subject repeats a phrase, a "1" is placed beside that category if at least one of that type intrusion was present. If the intrusion was not present, a "0" is placed beside the category. Each subject's responses were tape recorded so that accurate scoring could be achieved. The categories are not mutually exclusive; any one of them can occur in any combination with any of the other six. The symbolic representation and verbal description of each of the categories is as follows (Note: each phrase contains four single-syllable words, e.g., "Bud Buggs bleeds blood"):

$\underline{>4}$: This means that more than four words were spoken; for example, "Bud Bluggs Buggs bleeds blood".

P(ause): This means that the repetition was interrupted momentarily by a short pause between words; for example, "Bud Buggs...bleeds blood".

B(lock): This means that speech was blocked for at least a second between words; for example, "Bud Buggs.......... bleeds blood".

S(lur): This means that the sound of a word was strung out or prolonged; for example, "Bud Buggs bleeeeeds blood".

I(ntrusion): This means that some extraneous sound was produced; for example, "Bud Buggs (laughs) bleeds blood".

P(ause)/b(etween): This means that a pause was inserted between repetitions; for example, "Bud Buggs bleeds blood.........Bud Buggs bleeds blood".

M(iscellaneous): This means that a miscellaneous disruption, deviation, or intrusion not covered by any other category was present.

Predictions when the Dependent Variable is Speech Response Time

Introduction

Any of the three models of the possible effects of variable speech motor activity on the temporal relationship between implicit and explicit speech presented earlier can be factorially related to any of the five speech control hypotheses. However, a given implicit speech model may or may not be consistent with all of the speech motor control hypotheses. On what basis can a judgment as to the relationship between a given implicit speech model and a given speech control hypothesis be made? The present author feels that the most sound judgment can be made by matching the prevailing control tendency of a given hypothesis with the characteristic functional disposition of implicit and explicit speech. For example, if the prevailing control tendency of a given hypothesis matches the characteristic functional disposition of implicit speech (but not explicit speech), then it is likely that the primary function of that control unit is to control implicit speech; only secondarily would it control explicit speech (see Figure 2, page 20). Conversely, if the tendency of a given hypothesis matches the disposition of explicit speech (but not implicit speech), then it is likely that the control unit's primary function is to control explicit speech (see Figure 3, page 21). If the tendency matches both the implicit and explicit dispositions, they are probably controlled alike (see Figure 1, page 18). The prevailing control tendencies of the hypotheses were listed in Table I, page 15 (e.g., the tendency of the open loop is to control rapid speech).

The characteristic functional dispositions of implicit and explicit speech can be described as follows. Implicit speech is not used for communication (Vygotsky, 1962; Locke, 1970; Sokolov, 1972). Therefore, implicit speech need not be detailed, precise, or even totally fluent and accurate. It is not uncommon to hear a person exclaim that his thoughts are jumbled, and in many a mental hospital reside persons described as having disordered thoughts. Quickness is usually the characteristic demanded of implicit speech, e.g., a man being quick of mind. On the other hand, since explicit speech is used for communication quickness is not usually desirable. Accuracy is usually the characteristic demanded of explicit speech, e.g., could you speak more slowly and clearly? In summary, the characteristic functional disposition of implicit speech appears to be speed, and the characteristic functional disposition of explicit speech appears to be accuracy.

<u>Grammatical-Lexical Hypothesis</u>

The Grammatical-Lexical hypothesis assumes that speech motor control is designed to promote both rapid and accurate speech production. This tendency matches the characteristic functional disposition of both implicit and explicit speech. This match coincides with the "implicit and explicit speech as an end process" model (see Figure 1). Therefore, no difference is expected between implicit and explicit speech response times.

The major control problem concerning this hypothesis is the complexity of the relationship between phonemic specifications and articulatory motor movements. A Fast speech set should put pressure on the system to ignore complexities in order to have more rapid speech. A Perfect speech set should put pressure on the system to carefully control the complexities. T phrase processing should be more complex than NT processing. All of this means that the Fast set should produce faster speech than the Perfect set; T phrases should be spoken more slowly than NT phrases; the T/NT response time difference should be less under the Fast set than under the Perfect set.

A partial summary of these predictions is as follows.

For both Fast and Perfect conditions:

1. T phrases should have significantly slower speech response times than NT phrases.
2. There should be no significant difference between implicit and explicit speech response times.

Open Loop Hypothesis

The open loop hypothesis assumes that speech motor control is designed to promote rapid speech production. This tendency matches the characteristic functional disposition of implicit speech. This match coincides with the "implicit speech as in-process speech motor activity" model (see Figure 2). Therefore, it is predicted that explicit speech response time will be slightly slower than implicit speech response time.

The major control problem conterning this hypothesis is preprogramming. The greater the preprogramming the faster the speech. However, greater preprogramming also means less tight motor control and consequently a greater likelihood of disfluent or inaccurate speech.

Because of these consequences of variable preprogramming, the following relationships should hold. A Fast speech set should put pressure on the system to preprogram as much as possible. A Perfect speech set should put pressure on the system to preprogram as little as is practical. T phrases should have less preprogramming than NT phrases. All of this means the following. The Fast set should produce faster speech than the Perfect set. Since the open loop is designed for rapid speech control, the Fast set should predominate over the T phrase difficulty; therefore, under the Fast set T and NT phrases should be spoken at the same rate. Under the Perfect set, T phrases should be spoken more slowly than NT phrases.

A partial summary of these predictions is as follows.

Fast condition:

1. There should be no significant difference in speech response time between T and NT phrases.
2. Explicit phrases should have a slightly slower speech response time than implicit phrases.

Perfect condition:

1. T phrases should have significantly slower speech response times than NT phrases.
2. Explicit speech should be slightly slower than implicit speech.

Closed Loop Hypothesis

The closed loop hypothesis assumes that speech motor control is designed to promote accurate speech production. This tendency matches the characteristic functional disposition of explicit speech. This match coincides with the "implicit speech as neural speech command activity" model (see Figure 3). Therefore it is predicted that explicit speech response time will be slower than implicit speech response time.

The major control problem concerning this hypothesis is the critical difference criterion between target input and speech output--the less tight the criterion, the more rapid the speech processing but the less accurate the speech. Because of these consequences of variable criterion control, the following relationships should hold. A Fast speech set should put pressure on the system to make a change in the critical difference criterion, namely, to make it less tight. A Perfect set should put pressure on the system to change the criterion so that it is tighter. T phrases should have a tighter criterion than NT phrases. All of this means the following. The Fast set should produce faster speech than the Perfect set. T phrases should be spoken more slowly than NT phrases. The T/NT response time difference should be less under the Fast set than under the Perfect set.

A partial summary of the predictions is as follows.

For both Fast and Perfect conditions:

1. There should be no significant difference between T implicit and NT implicit speech response times.
2. T explicit phrases should have significantly slower speech response times than T implicit and NT implicit and explicit phrases.
3. NT explicit phrases should have slightly slower speech response times than T implicit and NT implicit phrases.

Open Loop-Closed Loop Hypothesis

The open loop-closed loop hypothesis assumes that speech motor control is designed to promote both rapid and accurate speech using two control loops to achieve this result. Since there are two control loops and two main functions of speech to control (rapidity and accuracy), the most efficient and parsimonious use of the two loops would seem to be if one loop controlled one speech function, and the second loop controlled the other speech function. Assuming this reasoning is valid, then the open loop would control rapid speech and the closed loop would control accurate speech.

This means that the characteristic functional dispositions of both implicit and explicit speech are matched. Since MacNeilage (1970) has speculated that the open loop operates prior to the closed loop, implicit speech should function prior to explicit speech. This relationship coincides with the "implicit speech as in-process speech motor activity" model (see Figure 2). Therefore, it is predicted that explicit speech response time will be slower than implicit speech response time.

There are two major control problems concerning this hypothesis: (1) for the open loop the problem is preprogramming; (2) for the closed loop the problem is the critical difference criterion. The consequences of variable preprogramming control and of variable criterion control should be exactly the same as described earlier for each individual loop, therefore the following relationships should hold. A Fast set should put pressure on the system to both preprogram as much as possible and also to make the critical difference criterion more flexible. A Perfect set should put pressure on the system to both program as little as possible and also to make the criterion less flexible. T phrases should have both less preprogramming and also a less flexible criterion than NT phrases. All of this implies that the Fast set should produce faster speech than the Perfect set. Under the Fast set, since implicit speech is under open loop control, T and NT phrases should be spoken at the same rate implicitly (see open loop predictions). Other than the forementioned exception, T phrases should be spoken more slowly than NT phrases; and the T/NT response time difference should be less under the Fast set than under the Perfect set.

A partial summary of the predictions is as follows.

Fast condition:

1. There should be no significant difference between T implicit and NT implicit speech response times.
2. T explicit phrases should have significantly slower speech response times than T implicit and NT implicit and explicit phrases.
3. NT explicit phrases should have significantly slower speech response times than T implicit and NT implicit phrases.

Perfect condition:

1. T explicit phrases should have significantly slower speech response times than T implicit, NT implicit, and NT explicit phrases.
2. T implicit phrases should have significantly slower speech response times than NT implicit phrases.

3. NT explicit phrases should have slightly slower speech response times than NT implicit phrases.

Closed Loop-Closed Loop Hypothesis

The closed loop-closed loop hypothesis assumes that speech motor control is designed to promote accurate speech production using two control loops to achieve this result. Since there are two control loops and only one main speech function to control (accuracy), the most efficient and parsimonious use of the two loops would seem to be if each controlled a different aspect of motor functioning (e.g., one loop controlled gross motor functioning and the other loop controlled fine motor functioning). Assuming this reasoning is valid and assuming further that gross motor functioning is more rapid than fine motor functioning, then one closed loop would control relatively rapid speech functioning and the other closed loop would control very accurate speech functioning. This means that the characteristic functional dispositions of both implicit and explicit speech are matched at least to some extent. Since MacNeilage (1970) has noted that gross motor functioning is generally prior to fine motor functioning, then implicit speech should function prior to explicit speech. This relationship coincides with the "implicit speech as in-process speech motor activity" model (see Figure 2). Therefore, it is predicted that explicit speech response time will be slower than implicit speech response time.

The major control problem concerning this hypothesis is the critical difference criteria for the two loops. Within each loop control unit, the consequence of variable criterion control should be exactly the same as described earlier for the single closed loop.

Therefore, the following relationships should hold. A Fast set should put pressure on the system to make the critical difference criterion more flexible for both loops. A Perfect set should put pressure on the system to make both criteria less flexible. T phrases should have less flexible criteria than NT phrases. All of this implies that the Fast set should produce faster speech than the Perfect set. T phrases should be spoken more slowly than NT phrases. The T/NT response time difference should be less under the Fast set than under the Perfect set. Since implicit speech is processed before explicit speech, the response time difference between implicit and explicit speech is expected to increase as a function of increasing difficulty of processed material.

A partial summary of the prediction is as follows.

For both Fast and Perfect conditions:

1. T explicit phrases should have significantly slower speech response times than T implicit, NT implicit, and NT explicit phrases.
2. T implicit phrases should have significantly slower speech response times than NT implicit and NT explicit phrases.
3. NT explicit phrases should have slightly slower speech response times than NT implicit phrases.

CHAPTER III

METHODS

Subjects

Thirty Oklahoma State University lower division undergraduate volunteer subjects were used, 15 in each of two between-subjects conditions: (a) Fast, and (b) Perfect. Subjects received extra course credit for their participation.

Stimuli

The stimuli consisted of three twister (T) phrases and three non-twister (NT) phrases. The T phrases were as follows: (1) Sue ships slip sheets, (2) Phil Phipps flips pits, (3) Bud Buggs bleeds blood. The corresponding NT control phrases were as follows: (1) John gets four cards, (2) Bob Finn hits nails, (3) Lynn Hall reads books.

Apparatus

For a given trial each of the phrases was to be repeated exactly 20 times implicitly or 20 times explicitly. Monitoring of the repetitions was achieved by having the subjects depress a microswitch every time they spoke the last word of the phrase. The microswitch was connected to a Hunter Klockounter so that the counter advanced one unit every time the switch was depressed.

The response time for the 20 repetitions was obtained as follows.

The experimenter would signal the subject to start the repetitions; at the same time the experimenter activated a Standard Electric clock. The subject depressed the microswitch every time he spoke the last word of the phrase; when the Klockcounter reached 20, the experimenter pressed a switch which stopped the timer, and signaled the subject to stop the repetitions.

A record of the speech errors was obtained by tape recording each session for later analysis.

Procedure

Each subject was randomly assigned to either the "Fast" or the "Perfect" condition. In addition, each subject performed the phrases according to a list which was randomized for both T/NT phrases and implicit/explicit speech. The subject was told that he would be speaking the phrases both aloud and silently, that each phrase should be continually repeated until he was told to stop, that he should depress the microswitch every time he spoke the last word of the phrase, and that he should begin repeating the phrase as soon as the experimenter said "Go".

The subjects in the "Fast" condition were instructed as follows:

> Don't think about errors---your job is to say the four words of the phrase as fast as possible no matter how many errors you make. Make an honest attempt to say the words, but if something other than what you wanted to say comes out it doesn't matter as long as you don't mumble or just make noise. Thus, the only restrictions are that words or word approximations are spoken (correct or not) and that the original word order of the phrase is attempted (whether the attempt is successful or not). Absolutely do not stop or hesitate. Say something, keep going. Absolutely do not correct or repeat missaid words or phrases. Do not slow

> down for better accuracy. Do not pace yourself for better accuracy. Do not put any contingency on yourself that would make you more accurate at the expense of slowing you down.

The subjects in the "Perfect" condition were instructed as follows:

> I want you to say the phrase as accurately as possible--no errors. However, I do not want you to go so slowly that you have exaggerated accuracy. The best method is to imagine a rate that you think will just get by without errors and then go just a little slower than that. If you make an error, don't worry about it except to slow down your speaking rate so that it does not happen again. I will stop you if you make too many errors and you'll have to start again. Absolutely do not correct an error if you make one; do not hesitate at any time; do not stop at any time unless I tell you to.

The subject then practiced the procedure with a practice T phrase--"Fred Flute picks fruit". If he did not follow the procedure correctly he was told what he was doing wrong and how to proceed correctly. The subject practiced until he was able to follow the procedure correctly for five repetitions. The subjects was then shown a card containing the six test phrases and was asked to read each one out loud. The card was then taken away and the subject was asked if he had any questions.

Then the experimenter started the tape recorder. Next, the subject was told which phrase he was to repeat and whether he was to repeat it aloud (explicitly) or silently (implicitly). If the subject had difficulty remembering the phrase, he was shown the card again. When the subject was ready, the experimenter said "Go" and started the timer. When the counter reached 20, the experimenter stopped the timer and told the subject to stop. The experimenter then recorded the response time, reset the timer and the counter and proceeded with the next trial. This continued until all six phrases were repeated 20 times both aloud and silently. The tape recorder was then turned off and the subject was debriefed.

CHAPTER IV

RESULTS

Each subject's mean response time score for the three phrases (20 repetitions per phrase) for each of the conditions (T implicit, T explicit, NT implicit, NT explicit) was obtained. Then the mean response times per subject for 20 repetitions of a phrase for all conditions was calculated. These means are presented in Table II. Figure 4 presents this same information graphically.

TABLE II

MEAN RESPONSE TIMES PER SUBJECT PER 20 REPETITIONS OF A PHRASE (IN SECONDS)

Speaking Set	Difficulty Factor	Speaking Mode	Response Time/ 20 Repetitions
Fast	Twister	Explicit	38.43
Fast	Twister	Implicit	28.03
Fast	Non-twister	Explicit	23.73
Fast	Non-twister	Implicit	21.29
Perfect	Twister	Explicit	52.39
Perfect	Twister	Implicit	43.28
Perfect	Non-twister	Explicit	33.00
Perfect	Non-twister	Implicit	29.22

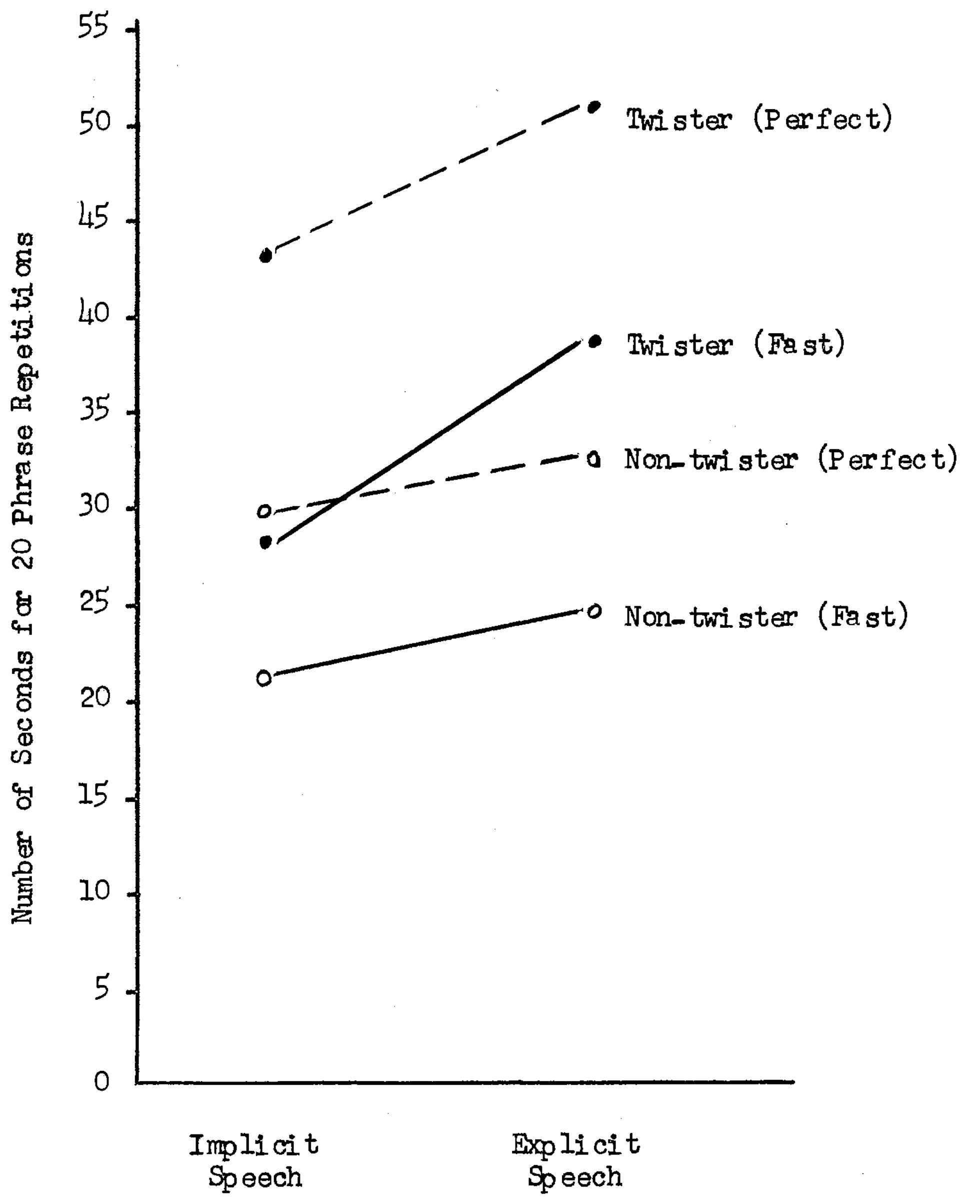

Figure 4. Response Time (Seconds/20 Phrase Repetitions) as a Function of Speaking Set, Level of Difficulty, and Speaking Mode

A split-plot 2x2x2 analysis of variance was performed with each subject's mean speech response time per 20 repetitions as the dependent variable. The results of the analysis showed the following. Response times are greater in the Perfect condition than in the Fast condition: $\underline{F}$ (1,28) = 31.66, $\underline{p}$ <.01. Response times are greater for the T phrases than for the NT phrases: $\underline{F}$ (1,28) = 236.61, $\underline{p}$ <.01. Response times are greater for explicit speech than for implicit speech: $\underline{F}$ (1,28) = 38.81, $\underline{p}$ <.01. The Fast-Perfect x T-NT interaction is significant: $\underline{F}$ (1,28) = 11.34, $\underline{p}$ <.01. The Fast-Perfect x implicit-explicit interaction is not significant: $\underline{F}$ (1,28) = .00, $\underline{p}$ >.10. The T-NT x implicit-explicit interaction is significant: $\underline{F}$ (1,28) = 29.63, $\underline{p}$ <.01. The Fast-Perfect x T-NT x implicit-explicit interaction is not significant: $\underline{F}$ (1,28) = 1.17, $\underline{p}$ >.10.

Post hoc Neuman-Keuls tests were performed to compare significance among means. The results are shown in Table III.

Intrusions are now considered. It should be noted that intrusions could only be counted in the explicit condition. Table IV lists the mean number of intrusions/20 repetitions based on the means of the three phrases for all 15 subjects/condition.

The reliability of the intrusion scoring was checked by having a second rater independently score intrusions on a sample of four subjects. For training, this second rater was shown the intrusion definitions as gives on page 23, along with several examples of each type of intrusion. Reliability was computed as percent agreement per category per repetition per phrase. The results were then combined into an overall percent agreement per category for all subjects in all conditions. This combined data is presented in Table V.

TABLE III

NEWMAN-KEULS MEAN COMPARISONS (MEANS BASED ON RESPONSE TIMES PER 20 REPETITIONS OF A PHRASE [IN SECONDS])

Mean Comparison	Critical Value $\alpha = .05$	Critical Value $\alpha = .01$	Difference Between Means
PTE/FTE	6.49	8.74	13.96**
PNTE/FNTE	6.49	8.74	9.26**
PTI/FTI	6.49	8.74	15.25**
PNTI/FNTI	6.49	8.74	7.92*
PTE/PNTE	3.12	4.20	19.39**
FTE/FNTE	3.12	4.20	14.70**
PTI/PNTI	3.12	4.20	14.07**
FTI/FNTI	3.12	4.20	6.74**
PTE/PTI	3.46	4.66	9.10**
PNTE/PNTI	3.46	4.66	3.78*
FTE/FTI	3.46	4.66	10.40**
FNTE/FNTI	3.46	4.66	2.44

F = Fast
P = Perfect
T = Twister
NT = Non-twister
E = Explicit
I = Implicit
* $p < .05$
** $p < .01$

TABLE IV

MEAN NUMBER OF INTRUSIONS PER 20 REPETITIONS OF A PHRASE (MAXIMUM 20 INTRUSIONS PER CATEGORY) FOR THE FOUR COMBINATIONS OF EXPLICIT CONDITIONS

Category	Perfect Twister	Perfect Non-twister	Fast Twister	Fast Non-twister
>4	2.20	.22	2.18	.67
P	5.27	.40	4.78	.87
B	.80	.00	.40	.00
S	3.16	.07	2.36	.56
I	.62	.09	1.78	.16
P/B	2.24	.69	2.51	1.16
M	.00	.00	.04	.16

TABLE V

PERCENT INTER-OBSERVER AGREEMENT PER CATEGORY FOR EACH REPETITION OF EACH PHRASE AVERAGED (MEAN) OVER A SAMPLE OF FOUR SUBJECTS ON ALL EXPLICIT CONDITIONS

			Categories			
>4	P	B	S	I	P/B	M
95.84	91.88	99.17	90.00	98.75	94.59	100

Because the number of intrusions was surprisingly large, intrusions were analysed in an attempt to discover the relationship between intrusions and the various experimental conditions. Therefore, the raw intrusion data was collapsed from seven categories into two mutually exclusive categories--intrusion and no intrusions. Next the number of repetitions/20 repetitions containing no intrusions versus number of repetitions/20 repetitions containing at least one intrusion were calculated as means. These data are presented in Table VI.

TABLE VI

MEAN NUMBER OF INTRUSION AND NO INTRUSION SCORES PER 20 REPETITIONS BASED ON THE FOUR COMBINATIONS OF EXPLICIT CONDITIONS

	Perfect Twister	Perfect Non-twister	Fast Twister	Fast Non-twister
Category	Mean	Mean	Mean	Mean
Intrusions	10.40	1.64	15.31	3.91
No Intrusions	9.60	18.36	4.69	16.09

Intrusions were analysed by a split-plot 2x2 analysis of variance. The results showed the following. There were fewer intrusions in the Perfect condition than in the Fast condition: F (1,28) = 21.87, $p < .01$. There were fewer intrusion in the NT condition than in the T condition: F (1,28) = 327.38, p <.01. The Fast-Perfect x T-NT interaction was significant: F (1,28) = 5.64, p <.05.

Post hoc Neuman-Keuls tests were performed to compare significance among means. The results are shown in Table VII.

TABLE VII

MEAN COMPARISONS FOR INTRUSION SCORES BASED ON POST HOC NEUMAN-KEULS TESTS

Mean Comparison	Critical Value α = .05	α = .01	Difference Between Means
FT/FNT	1.61	2.17	11.40**
PT/PNT	1.61	2.17	8.76**
FT/PT	1.94	2.61	4.99**
FNT/PNT	1.94	2.61	2.27*

F = Fast
P = Perfect
T = Twister
NT = Non-twister
* p <.05
** p <.01

Because of the unexpected high number of intrusion errors and because of the possible relationship between intrusions and speech response time, it was decided to retime the tapes on a sample of intrusion-free speech. This was done as follows. Five subjects each from the Perfect and the Fast conditions were found to have at least five consecutive repetitions which were free from intrusions for all three phrases in both the T and the NT conditions. Response times for these phrases were obtained by timing the explicit condition. To obtain comparable times for the implicit condition, the original implicit response times (for 20 repetitions) were divided by four. The results are presented in Table VIII.

TABLE VIII

MEAN RESPONSE TIMES PER SUBJECT PER FIVE REPETITIONS OF A PHRASE (IN SECONDS)

Speaking Set	Difficulty Factor	Speaking Mode	Response Time/ Five Repetitions
Fast	Twister	Explicit	7.84
Fast	Twister	Implicit	6.90
Fast	Non-twister	Explicit	5.85
Fast	Non-twister	Implicit	5.40
Perfect	Twister	Explicit	11.09
Perfect	Twister	Implicit	10.75
Perfect	Non-twister	Explicit	8.14
Perfect	Non-twister	Implicit	7.27

A split-plot 2x2x2 analysis of variance was performed with each subject's mean speech response time per five repetitions as the dependent variable. The results of the analysis showed the following. Response times are greater in the Perfect condition than in the Fast condition: $\underline{F}$ (1,8) = 4.83, $\underline{p} < .10$. Response times are greater for the T phrases than for the NT phrases: $\underline{F}$ (1,8) = 82.60, $\underline{p} < .01$.

Response times for explicit speech and implicit speech are not significantly different: $\underline{F}$ (1,8) = 2.06, $\underline{p} > .05$. The Fast-Perfect x T-NT interaction is significant: $\underline{F}$ (1,8) = 7.28, $\underline{p} < .05$. The Fast-Perfect x implicit-explicit interaction is not significant: $\underline{F}$ (1,8) = .01, $\underline{p} > .05$. The T-NT x implicit-explicit interaction is not significant: $\underline{F}$ (1,8) = .003, $\underline{p} > .05$. The Fast-Perfect x T-NT x implicit-explicit interaction is not significant: $\underline{F}$ (1,8) = 2.21, $\underline{p} > .05$. These results are presented graphically in Figure 5.

Post hoc Neuman-Keuls tests were performed to compare significance among means. The results are shown in Table IX.

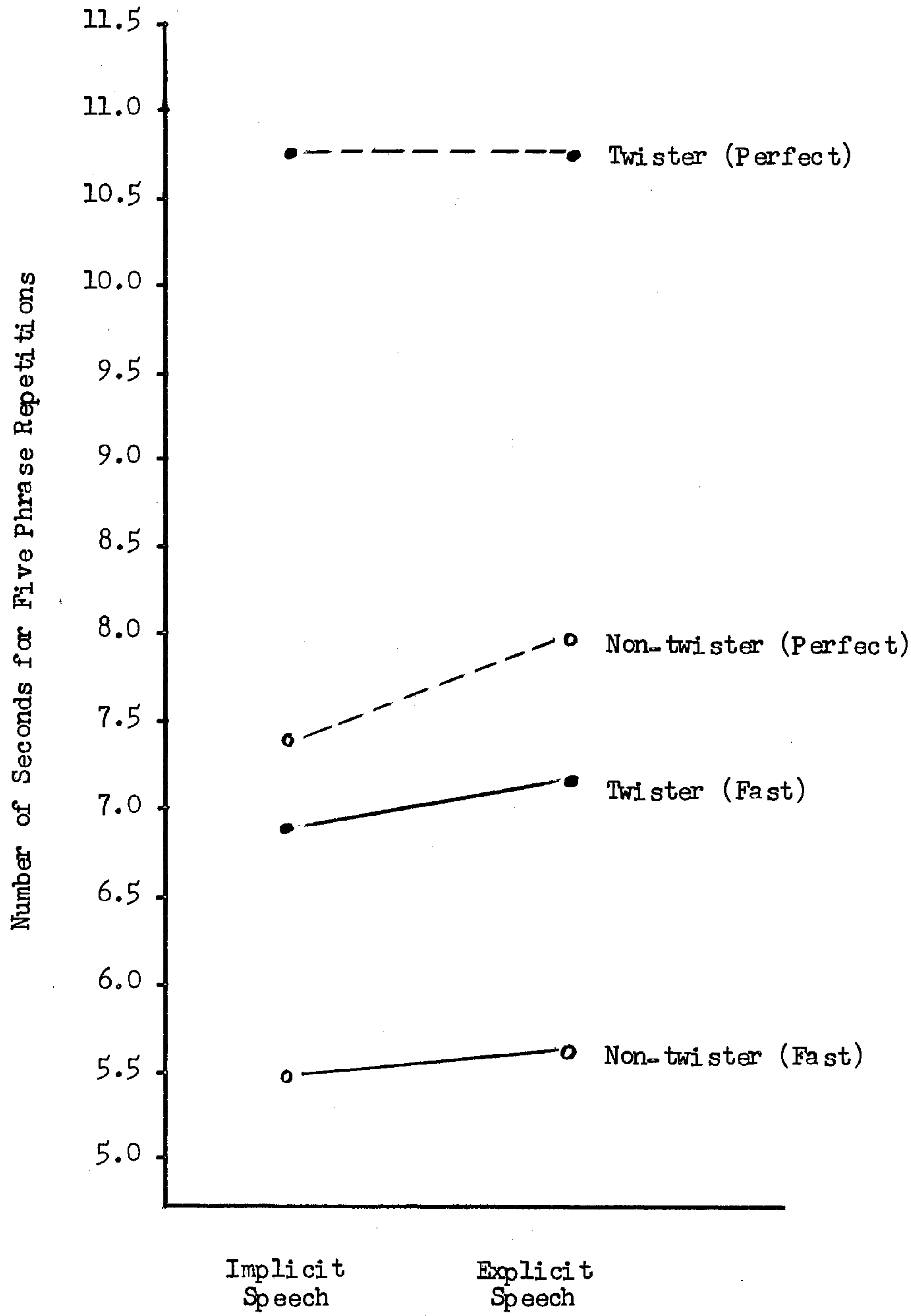

Figure 5. Response Time (Seconds/Five Phrase Repetitions) as a Function of Speaking Set, Level of Difficulty, and Speaking Mode

TABLE IX

NEWMAN-KEULS MEAN COMPARISONS OF RT FOR INTRUSION-FREE DATA (MEANS BASED ON RESPONSE TIMES PER FIVE REPETITIONS OF A PHRASE [IN SECONDS])

Mean Comparison	Critical Value $\alpha = .05$	Difference Between Means
PTE/FTE	4.13	3.24
PNTE/FNTE	4.13	2.29
PTI/FTI	4.13	3.85
PNTI/FNTI	4.13	1.87
PTE/PNTE	1.02	2.95*
FTE/FNTE	1.02	1.99*
PTI/PNTI	1.02	3.48*
FTI/FNTI	1.02	1.50*
PTE/PTI	1.54	.34
PNTE/PNTI	1.54	.87
FTE/FTI	1.54	.94
FNTE/FNTI	1.54	.45

F = Fast
P = Perfect
T = Twister
NT = Non-twister
E = Explicit
I = Implicit
* $p < .05$

CHAPTER V

DISCUSSION

The Speech Production Hypotheses

The response time data are supportive of the closed loop-closed loop hypothesis; the results fit the predictions quite well. However, it should be remembered that the response time predictions were made under the assumption that intrusions should not occur. The intrusion data indicate that a high percentage of intrusions did occur, and it is possible that these intrusions, rather than closed loop-closed loop mechanisms, were responsible for the patterning of the response times.

In an attempt to discern the role that these intrusions played in the patterning of the response times, samples of intrusion-free speech were timed and analyzed. The results of this procedure were presented in Tables VIII and IX. As can be seen, the resultant response time comparisons differ dramatically from those of the original data. The intrusion-free comparisons indicate that the twister phrases were spoken more slowly than the non-twister phrases, and that explicit speech had nearly the same rate of response as implicit speech for all experimental conditions. These results support the Grammatical-Lexical hypothesis.

Because of the difference between the intrusion-free results and the original results, the question arises as to how much confidence can be placed in the intrusion-free data. The converging information that can be brought to bear on this issue all tends to support the validity of the intrusion-free results. First, consider the response time information. (A) All previous work comparing implicit and explicit speech response times has found the times to be equal (Landauer, 1962; Weber & Bach, 1969; Weber & Castleman, 1970). In addition, Landauer (1962) noted that implicit speech sets take slightly less time at first than explicit speech sets. Every one of the comparisons presented in Tables VIII and IX conforms to both of these previous findings. (B) Since the consistency of the intrusion-free explicit/implicit comparisons transcends the Fast/Perfect between-subjects groups, this indicates that the results are probably not due to an unusual sample of subjects.

Second, consider that intrusion-free explicit speech could only be consistently equal in response time to the implicit speech if the implicit speech is also intrusion-free. What information is there that the implicit speech counterparts of the intrusion-filled explicit speech do not, in themselves, contain many errors? Consider the following. (A) The experimenter's experience in observing the button pushing by subjects during the repetition of phrases was that button pushing during implicit speech was always smooth and rhythmic while button pushing during dysfluent episodes of explicit speech was not smooth or rhythmic. (B) The experimental protocols of the subjects' comments about the experiment indicate that even though the subjects were questioned about the nature of their implicit speech, not one indicated any occurance that resembled an instance of an intrusion.

To explain the discrepancy between the original findings and the intrusion-free findings, consider the following. Mysak (1966) has speculated on the existence of multiloop control of speech production. In part, his proposed system works as follows: "the individual continually scrutinizes speech content output, compares it with his thoughts, and makes appropriate adjustments when necessary" (p. 18). He has further speculated that this multiloop control has at least two levels of activity, a higher "thinking level" devoted to thought and speech content monitoring, and a lower more automatic "doing level" devoted to articulatory motor process control.

In reviewing the tapes of the subjects' performances, it was noticed that ≥ 4 intrusions occurred mainly in the form of correcting misspoken words, and that intrusions such as laughs and gasps occurred only during dysfluent episodes. Both of these intrusion behaviors appear to be attempts to make "appropriate adjustments" because "thought" and "speech output" did not match.

Therefore, it appears that the original response time data might represent the combination of Grammatical-Lexical motor control process response time, and a closed loop monitoring process response time. This closed loop monitoring process could be concerned with neural speech command activity. Since this closed loop would only make adjustments of explicit speech, only explicit speech would show an increase in response time when closed loop response time was combined with the motor control response time. Furthermore, explicit speech for T phrases (which are more likely to get out of synchronization than NT phrases) would be the most vulnerable to adjustments and therefore an increase in response time. Therefore, the combination of closed loop monitoring response time and motor control response time would be greatest for explicit T phrases. This is what the original results show.

Other Considerations

Intrusions

Intrusions were found to increase both with rapid speech and with difficult speech. These effects interact in that the Fast twister phrases had more intrusions than any other condition. Twisters seem to be more effective than rapid speech in producing intrusions as evidenced by the fact that the Perfect twister phrases produced more intrusions than the Fast non-twister phrases. This appears to be an example of a general finding that type of material spoken has a stronger effect on speaking than does instructional orders on how to speak. Subject corrections of mistakes were not eliminated by the instructional admonition to subjects not to correct. In addition, there were many intrusions under the Perfect instructional set. Both of these findings give some evidence that the influence of this monitoring loop is fairly strong and possibly unavoidable. Also, the types of intrusions found--for example P/b's, B's, and P's--indicate that the monitoring loop might be able to intercede upon the ongoing speech process at any point in the explicit performance of that process.

Speaking Sets

The speaking sets (Fast/Perfect) did influence subjects' speech; subjects spoke more rapidly under the Fast set than under the Perfect set. Also, subjects spoke more accurately under the Perfect set than under the Fast set. The Perfect set, however, did not produce intrusion-free speech. This again argues that the type of spoken material has a stronger influence on speech than instructional orders.

Difficulty Factors

Twisters were spoken more slowly than non-tristers under all conditions. This indicates that some articulatory switches are more difficult than others.

Speaking Mode

The intrusion-free data indicates that implicit and explicit speech are spoken at approximately the same rate.

Suggestions for Future Research

The following are suggestions for future research. The effects of an auditory mask on the nature and frequency of intrusions could be studied. The present experiment would be replicated with an additional condition in which subjects listen to white noise over headphones during their performance (repetition of the test phrases). The results should give some indication of the relationship between intrusions into speech and self-monitoring of speech.

The effects of practice and pretraining on the nature and frequency of intrusions could be studied. An attempt could be made to eliminate intrusions by massive amounts of pretraining. In essence, what would be examined would be the feasibility and possibility of making a twister phrases in to a non-twister phrase. A possible extension of this idea would be to introduce anomalous twisters (e.g., Blug blug glub bug) into the phrase pool and try the massive pretraining with them also. The results should give some indication of the relationship between intrusions into speech and the difficulty of the spoken material. Also, the results should give some indication as to whether the speech system can completely adapt to difficult and anomolous material, thus eliminating the difference in response times found in the present experiment.

A longitudinal or comparative study could be performed to test the hypothesis that children have different speech motor control than do adults. One speculation is that children have open loop control due to their need for flexibility and learning while adults have Grammatical-Lexical control because of their need for efficiency and accuracy.

Preprogramming of speech or the lack thereof could be studied by presenting twister and non-twister phrases visually and recording the response latency from the time of onset of the visual presentation to the utterance of the first sound. The assumption tested would be that the greater the need for preprogramming, the greater the response latency. Difficulty of the phrase and length of the phrases could be varied as independent variables. Also instructional sets, e.g., "speak fast" or "speak perfectly", could be used.

Finally, it seems that tongue twister difficulty is an important parameter for future research in the area of speech production and control. More information is needed as to the nature and causation of tongue twisters.

Summary

In summary, intrusion-free data of the present experiment supports the Grammatical-Lexical hypothesis of speech control at the articulatory level. It supports the "implicit and explicit speech as end product" conception of their relationship. It also supports the assumption that T phrases are harder to control than NT phrases. Finally, both the intrusion score data and the original response time data give evidence for a closed loop monitor that controls the synchronization of explicit speech and the neural speech command activity, and appears to have a powerful intrusive effect on speech production.

BIBLIOGRAPHY

Fromkin, V.A. The non-anomalous nature of anomalous utterances. Language, 1971, 47, 27-52.

Hanfmann, E., & Vakar, G. (Translators) Vygotsky's Thought and Language. Cambridge: M.I.T. Press, 1962.

Hockett, C. F. Where the tongue slips, there slip I. In To Honor Roman Jakobson (Janua Linguarum, Series Major, 32). The Hague: Mouton, 1967, 2, 910-936.

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole Pub. Co., 1968, 298-303.

Landauer, T. Rate of implicit speech. Perceptual and Motor Skills, 1965, 15, 646.

Lashley, K. S. The problem of serial order in behavior. In Brain and Behavior 2: Perception and Action. Pribram, K. E. (Ed.) Baltimore: Penguin Books, 1969, 515-540.

Liberman, A. M. Some results of research of speech perception. Journal of the Acoustical Society of America, 1957, 29, 117-123.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. Psychological Review, 1967, 74, 431-461.

Locke, J. L. Subvocal speech and speech. ASHA, 1970, 12, 7-14.

MacKay, D. G. The structure of words and syllables: evidence from errors in speech. Cognitive Psychology, 1972, 3, 210-227.

MacNeilage, P. F. Motor control of serial ordering of speech. Psychological Review, 1970, 77, 182-196.

McGuigan, F. J. (Ed.) Thinking: Studies in Covert Language Processes. New York: Apple-Century-Crofts, 1966.

Miller, G. A. The magic number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Mysak, E. D. Speech Pathology and Feedback Theory. Springfield, Illinois: Charles C. Thomas, 1966.

Novikova, L. A. Electrophysiological investigation of speech. In Thinking: Studies in Covert Language Processes. McGuigan, F. J. (Ed.) New York: Apple-Century-Crofts, 1966.

Sokolov, A. N. Inner Speech and Thought. New York: Plenum Press, 1972.

Sussman, H. M. What the tongue tells the brain. Psychological Bulletin, 1972, 77, 262-272.

Vygotsky, L. S. Thought and Language. Hanfmann, E., & Vakar, G. (Translators) Cambridge: M.I.T. Press, 1962.

Weber, R. J., & Bach, M. Visual and speech imagery. British Journal of Psychology, 1969, 2, 199-202.

Weber, R. J., & Castleman. J. The time it takes to imagine. Perception and Psychophysics, 1970, 8, 165-168.

Weiss, D. A. Cluttering. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.

VITA

Roger Walter Harnish

Candidate for the Degree of

Master of Science

Thesis: INVESTIGATION OF SPEECH PRODUCTION PROCESSES

Major Field: Psychology

Biographical:

Personal Data: Born in Brattleboro, Vermont, November 18, 1945, the son of Mr. and Mrs. Roger M. Harnish.

Education: Graduated from West High School, Rochester, New York, in June, 1963; received the Bachelor of Arts degree in Psychology from the University of Rochester, Rochester, New York, in June, 1970; completed requirements for the Master of Science degree at Oklahoma State University, Stillwater, Oklahoma, in December, 1973.

Professional Experience: Research Assistant, Department of Psychology, University of Rochester, 1970; Counselor, University Counseling Service, Oklahoma State University, 1972; Counselor, Psychological Guidance Center, Oklahoma State University, 1972; NDEA Fellow and Teaching Assistant, Oklahoma State University, 1972; Graduate Teaching Assistant, Department of Psychology, Oklahoma State University, 1973.