SPEECH ERRORS AND SEGMENT DURATION: AN INVESTIGATION OF WORD-INITIAL /sp, st, sk/-CLUSTERS UNDER CONDITIONS OF RAPID REPETITION

by

Verna Lynne Pyplacz

B.A., University of British Columbia, 1974

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Paediatrics

Division of Audiology and Speech Sciences

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
May, 1976

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Verna Lynne Pyplacz

Department of Paediatrics
Division of Audiology and Speech Sciences
The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 14th May 1976.

ABSTRACT

Speech errors, or "slips of the tongue", have been studied in attempts to understand the speech production process, to investigate phonological units and rules, and to provide insights into historical linguistic change. The present study examines speech errors and their relation to segment durations in word-initial /sp, st, sk/-clusters produced under rapid repetition conditions by six adult native speakers of English.

Fifty percent of the errors produced could be classified as repetition errors; these were examined for duration in the initial clusters, both error and corrected productions. General results following from analysis of the data were:
(1) Error clusters and their component segments were consistently longer in duration than their subsequent and immediate corrections.
(2) The clusters /sp/ and /sk/ are longer than /st/, which may be attributable to the faster moving, more highly innervated tongue tip musculature involved in the production of /s/ and /t/, compared with the heterorganic clusters.
(3) The stop consonant in a given cluster appears to determine the overall cluster duration, since the duration of /s/ remains fairly constant irrespective of context.

In light of the results, it was speculated that the excessive duration of the cluster (or of its component parts)

violated a timing constraint on the production of an utterance, necessitating recalibration and correction of the error. It was further inferred that feedback must be present in order for the system to recognize the duration error, to compare it with planned output, and finally to execute a correction.

Two types of feedback were considered necessary for the adequate functioning of a speech production model, which would also allow for speech perception: (a) continuous auditory feedback, which is supplemented by (b) intermittent proprioceptive feedback, both of which are used in perceiving input and manipulating output. Such a system provides a plausible account of speech error production as described in this study. The hypothesized variable servomonitor system advocated here (and in other studies) in general provides an efficient means for producing, monitoring and correcting speech production.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table | Page

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to express my gratitude to all those who gave of their time and energy in helping me complete this thesis. My thanks go especially to:

- Dr. John Gilbert -- for his time and patience given as my supervisor;
- Dr. Dale Kinkade -- for transcribing data, giving critical comments, and being kind enough to serve on my committee;
- Dr. John Delack -- who doesn't suffer fools gladly, for his time, effort and necessary critical comments;
- Ronnie Sizto -- for days (and nights) spent patiently running, debugging, rerunning and interpreting statistical programmes at the eleventh hour;
- Dr. Brenda Frazer -- for her statistical advice;
- Ingrid Jeffrey -- for her advice on analyzing oscillograms;
- My Subjects -- for their patience in repeating "tongue-twisters";
- My Family -- who are always there when I need them;
- My Friends in the Graduating Class of '76.

This thesis is dedicated
to the memory of my father,
Stan Pyplacz (1922 - 1975).

CHAPTER 1

INTRODUCTION

1.0 Introduction

A speech error can be, and has been, variously described as a "spoonerism" (after the Revd. William Spooner), a "portmanteau" word (coined by Lewis Carroll) or a "slip of the tongue" and can best be defined as "an unintentional linguistic innovation" (Sturtevant, 1947, p. 38). Speech errors are constrained by the grammar and phonology of a given language, and because of these constraints they are to a certain degree predictable and non-random (Fromkin, 1973, p. 113).

Authors such as Shakespeare, Rabelais and Lewis Carroll used speech errors in their works to achieve humourous ends. Freud believed that these disturbances of speech were "the result of complicated psychical influences, of elements outside the same word, sentence or sequence of spoken words" (Freud, 1924; cited in Fromkin, 1973, p. 110). Speech errors have also been studied in the hope that such would provide insight into historical linguistic change (Sturtevant, 1947), in attempts to understand the speech production process more fully (Boomer & Laver, 1968; Fromkin, 1968, 1971; Nooteboom, 1969), and to investigate the possible bases for certain phonological units and rules (Fromkin, 1968, 1971).

Few investigators have speculated as to the underlying cause of speech errors, although Meringer (1908) tried -- and failed -- to correlate error production with numerous variables, such as rate of speech and time of day. While speech errors have been used as a vehicle for investigating various speech processes, this temporary "breakdown and re-calibration" process in which the system is involved can be studied in its own right; i.e., an examination of the intrinsic structure of speech errors is logically prior to their use for other purposes.

The methodologies employed to collect speech error data, as well as the subsequent classifications of such data, have been many and diverse. The speech errors collected in the present study, for example, could be classified and accounted for almost entirely by the descriptions provided in Fairbanks and Guttman (1958); with regard specifically to the "repetition errors" found in the present study, researchers in the field of delayed auditory feedback (DAF) have encountered a similar phenomenon which they have labelled "artificial stutter" (cf. Lee, 1951). In the DAF literature, it has been noted that subjects require a "turn around time" (i.e., a delay in which a subject can produce a repetition after a first production), and this delay time could have a neurophysiological basis, such as that proposed by Kent and Moll (1975).

An examination of the literature pertaining to functional neuroanatomy and the neurophysiology of motor responses (e.g., Bowman, 1971; Abbs, 1973) suggests that this delay arises as a result of cortically controlled motor mechanisms (gamma and alpha motor systems); i.e., delay time might be accounted for by means of a gamma "delay" loop for feedback from the position and movement of the articulators (tongue, jaw, etc.) during speech.

1.1 Review of the Literature - Introduction

The literature relevant to this study will be discussed in four sections: (1) an outline of the functional neuroanatomy of speech, (2) feedback mechanisms, (3) speech errors, and (4) timing of speech. The fifth section will provide a discussion and summary of models of speech production in an attempt to synthesize information from the above-mentioned disciplines.

1.2 Functional Neuroanatomy of Speech

As has often been noted (e.g., MacNeilage, 1972, pp. 6-7), the importance of physiological mechanisms for speech is their interaction in the production of an acoustic output which has communicative significance. In this section a brief overview of some of the main areas of the brain with specific significance in generating speech will be discussed, as well as the central nervous system, the peripheral nervous system and the sensory and motor tracts, all of which make up the pathways for speech. The discussion will descend, anatomically speaking, from the cortex to the thalamus and on downward through the midbrain, pons, medulla, cerebellum, cranial nerves and spinal cord.

1.21 Central Nervous System

The cerebral cortex is to be regarded as the supreme manipulator of motor neuron impulses resulting in speech. This idea has been debated for several decades, culminating with the notion that the central nervous system (CNS) can be regarded "as a series of functional arcs in which subcortical centers are in a reciprocal relationship with cortical areas" (Berry & Eisenson, 1956, p. 45; cf. also Penfield & Roberts, 1959, p. 15). These arcs are claimed to interact with one another and not exclusively with the cortex. This view of the CNS as an input-output reflex arc is now, however, somewhat outmoded: Pribram (1971) has described the presence of feedback and feedforward mechanisms of the CNS which control receptor functions and has hypothesized a functional "Test-Operate-Test-Exit" servomechanism, which matches input against the output target. Whichever system is at work in the CNS, the result is a complex interactive process: speech.

Auditory perception, which is used for learning and maintaining speech, is believed to be found in the auditory reception area (Brodmann's Area 22). It is within this area that an individual may perceive sounds but not decode their meaning, this latter function being accomplished in Wernicke's area (Area 41-42):

> "In the auditosensory area [Area 22] auditory impressions reach consciousness as sounds, and their loudness, quality and pitch can be differentiated. The direction from which the sound comes and its character, whether rhythmical or arhythmical, are also determined by this part of the cortex. The significance and the source of the sound, however, require the adjoining auditopsychic area for their elucidation .... In this area [Areas 41-42] auditory impressions receive their interpretation and can be differentiated from one another, as regards their probable source and origin, by association with past experience." (Johnston & Whillis, 1954, p. 1037)

It has been demonstrated in recent years that loudness, quality and pitch can also be differentiated sub-cortically.

In the parietal lobe are Areas 1-3 which make possible awareness of touch, pressure, temperature and muscle movement. An awareness of tongue movements in articulation may be projected from here to speech areas and "may be one of the chief stimuli in provoking or continuing speech" (Berry & Eisenson, 1956, p. 56).

The motor projection area for voluntary movement (Area 4) has a large percentage devoted to phonatory and articulatory movement which sends impulses via this pyramidal tract to the muscles of the jaw, lips, tongue, larynx and pharynx.

Area 6, the extrapyramidal area, produces refinement in motor behaviour such as the sequencing of vocal fold adduction, resonance and articulation, or the qualities of intonation and rhythm.

Area 44, Broca's area, is where fibres from other areas concerned with the speech process synapse and then proceed to the motor projection areas for the muscles of speech. Other areas which may be similar in function to Area 44 are Areas 7A, 7B and 7C, the last of which is concerned with the thalamus and with integration of emotional expression into speech.

Areas 8-11, the frontal ideational association areas, are called upon to integrate past experience, abstract thinking, reasoning and ideas into speech.

The striate bodies, comprised of the caudate, lenticular and amygdaloid nuclei, together with the cortex and the thalamus, probably act as one unit or arc (cf. Penfield & Rasmussen, 1950, pp. 106-107). The caudate and lenticular nuclei belong to the extrapyramidal system, and their axons run to motor nuclei of the brain stem concerned with innervation of the muscles of the tongue, face, larynx and pharynx.

The highest integrative mechanism for speech may, as Penfield and Rasmussen (1950, p. 219) suggest, "be situated in some cerebral area, such as the thalamus, and not in either cerebral cortex". All sensory tracts have a relay station in the thalamus. Sensory-emotional responses, the quality of the voice, facial expression and subtle body gestures, as well as conceptual patterns of form, size, quality, intensity and texture, are organized here for transmission to the cortex.

The midbrain contains the substantia nigra and red nuclei, which are part of the extrapyramidal system, and the cerebral peduncles, which contain the pyramidal tracts. The pyramidal tract is made up of the corticospinal fibres which run from the motor cortex to the spinal cord. This tract has control over the speech muscles of the head and neck through cranial nerves V, VII, IX, X and XII. The extrapyramidal system is chiefly made up of the structures other than the cortex which send impulses to the spinal cord, i.e., striate bodies, cerebellum, red nucleus, substantia nigra, etc. (cf. Netter, 1974). The substantia nigra and the red nuclei have two-way connections with the striate bodies, thalamus and premotor area of the cortex. The substantia nigra is believed to control the muscles of facial expression, and the red nuclei in association with the cerebellum control the gradation and timing of muscular contraction.

The pons, located just below the midbrain, contains sensory and motor pathways, as well as the reticular formation, which is linked with the cerebellum and striate bodies, making up part of the extrapyramidal system. The pneumotaxic centre of the pons is connected to the hypothalamus and stimulates exhalation and maintains respiratory rhythm for speech. The trigeminal sensory complex is the principle sensory nucleus of the trigeminal (Vth) nerve in the pons. The motor nuclei of the facial and trigeminal nerves in the pons innervate voluntary facial speech musculature and muscles of mastication, respectively.

The medulla contains the centres which control the respiratory and circulatory systems and also regulate rate and rhythm of breathing for speech. These centres respond to incoming sensory impulses from the diaphragm and from the aortic and carotid capillaries. The lower motor neurons of the medulla innervate muscles of the mouth, pharynx and larynx for speech production via specific cranial nerves (to be discussed below). The medulla contains the nucleus solitarius which receives afferent (sensory) fibres from the facial, vagus, and glossopharyngeal nerves. The hypoglossal nucleus of the medulla supplies innervation to tongue muscles. The nucleus ambiguus sends fibres through the glossopharyngeal, vagus, and spinal accessory nerves to supply muscles of the pharynx and larynx.

The cerebellum,part of the extrapyramidal system, is -- in addition to providing for fine motor coordination in general -- of importance in speech production, by elaborately controlling voluntary muscle movements, e.g., in the modulation of phonation and articulation.

The spinal cord conducts sensory impulses to higher centres, such as the cerebellum and the thalamus. It also mediates control of motor activities of the body below the face and neck (e.g., posture, movements and gesture). The spinal cord acts as an integrating centre for many reflex patterns.

1.22 Peripheral Nervous System

The cranial nerves directly associated with speech mechanisms are the trigeminal (V), facial (VII), glosso-pharyngeal (IX), vagus (X), accessory (XI) and hypoglossal (XII).

The trigeminal nerve (V), containing both sensory and motor fibres important to the articulatory movements of speech, transmits sensations of movement from the muscles of mastication of the jaw, sensations of touch, temperature and pain from the face, and voluntary motor impulses to the jaw.

The facial nerve (VII), as well as cranial nerves IX-XII, has motor fibres innervating the muscles of speech production mechanisms. The facial nerve itself supplies the striated muscles of the face, the stylohyoid muscle and the stapedius muscles.

The glossopharyngeal nerve (IX) innervates the stylopharyngeus muscle, which aids in velar closure. It likewise mediates proprioception of the posterior third of the tongue.

The vagus nerve (X), in conjunction with cranial nerves IX and XI, innervates the voluntary muscles of the pharynx and larynx involved in speech. Sensory impulses, transmitting proprioceptive information to the medulla, cerebellum and other parts of the extrapyramidal tract, can effect fine coordination, graded contraction and tonic control necessary for speech.

The spinal accessory nerve (XI) assists the vagus in motor control of the pharynx and larynx and innervates the trapezius and sternomastoid muscles necessary for speech breathing.

Cranial nerve XII, the hypoglossal, innervates the striated muscles of the tongue. These muscles concerned with voluntary movements of the tongue are the genioglossus, geniohyoid, hyoglossus, stylohyoid, styloglossus, and palatoglossus.

There are thirty-one pairs of spinal nerves, which transmit both sensory and motor information. The most important role of these nerves in speech production consists of sending motor (efferent) impulses to activate the muscles of breathing for speech.

1.23 Sensory Tracts

It is of importance to this discussion to remember that proprioception (muscle position, tension and movement) is responsible for the rate, force, direction and extent of voluntary movements. One of the largest systems of sensory tracts transmitting such information consists of the ventral and lateral spinothalamic tracts.

As an example of relays along a sensory tract, we will utilize the auditory pathway. The stimulus is transformed into an electrochemical impulse in the cochlea, and the first synapse is located in the cochlear nucleus of the medulla. The fibre tract ascends in the lateral lemniscus through the pons to the midbrain. There is another synapse in the inferior colliculus and a final one in the medial geniculate body of the thalamus, from which the information then passes to the auditory reception area (Area 22) and to other areas in the cortex.

While the auditory impulse is being transmitted, proprioceptive impulses from muscles in the tongue, lips and jaw (among others) are reporting shifts in position and tension which will be used in the production of speech (cf. Berry & Eisenson, 1956).

It is worth noting here the differences between efferent and afferent tracts. An afferent, or sensory, tract conducts an impulse from the periphery towards the cortex, while an efferent, or motor, tract conducts impulses from the CNS to efferent nerve endings in muscles.

1.24 Motor Systems

Two tracts comprising the pyramidal system, the corticospinal and corticobulbar tracts, originate in the motor and premotor areas of the cortex.

The corticobulbar tract is important for speech production, since it activates the muscles of the tongue, lips, jaw, pharynx and larynx.

The corticospinal tract, having cell bodies in the precentral gyrus (i.e., in the frontal lobe) of the cortex, makes its way via the internal capsule (a fibre tract) to the cerebral peduncles (i.e., a crossing of several fibre tracts) in the midbrain. This fibre tract decussates, or crosses, in the medulla and then passes into the spinal cord, synapsing with motor cells effecting voluntary muscle contraction. The corticobulbar tract follows the same route until it passes into the pons and medulla, synapsing with lower motor neurons of cranial nerves V and VII-XII. From here fibres of this tract continue on to innervate the muscles of speech.

The extrapyramidal system is also vital to the finely coordinated motor activity necessary for speech. Its organization is described by Grinker and Bucy (1949, p. 274) as follows:

> "All of these subcortical structures which are recipients of impulses from the precentral cortex have two principal projection systems. 1) They project to the lateral nucleus of the thalamus and thence back to the precentral region (cortex) and 2) they have a descending pathway down the spinal cord to the anterior horn cells. ... [The necessity of this system for speech activities is that it] controls, activates, and inhibits the associated musculature or protagonistic muscles which must be appropriately contracted. ... It controls the reflex innervation of the skeletal muscles to produce what is commonly known as tone." [Cited in Berry & Eisenson, 1956, p. 71]

As can be deduced from the above discussion of neuroanatomical structures and their functions, speech can be regarded as a very complex integrative process, encompassing not only motor systems for its production, but also sensory monitoring systems, which are to be discussed in more detail in the following section.

1.3 Feedback Mechanisms

1.30 Introduction

The goal of this section is an appreciation of the neurophysiological basis for peripheral proprioceptive/kinesthetic feedback systems in operation during the speech act. The presentation here is germane to the discussion of models of speech production in Section 1.6. Researchers, such as Abbs (1973), Hardy (1970) and Bowman (1971; cf. Smith's 1973 review of Bowman), have contributed most of the information discussed in this section. The auditory system as a feedback mechanism for speech is also discussed here, and the main body of this section's presentation comprises an introduction to the gamma, or spindle, motor system and its import for feedback.

1.31 <u>The Gamma/Spindle Motor System</u>

> "The gamma-loop can be considered as the gamma efferent fibers, the spindle fiber controlled by the efferent fibers, and the synaptic connections made by the spindle afferent with alpha motoneurons (after Smith, 1969)." (Abbs, 1973, p. 176)

A muscle spindle is a small cylindrical body to the main body of the muscle. It contains intrafusal and extrafusal muscle fibres. The motor innervation of the body of a muscle is carried out by alpha motoneurons in the cortex, while innervation of muscle spindles is carried out by gamma motoneurons. Sensory neurons in the muscle spindle form a monosynaptic reflex arc with alpha motoneurons.

It is instructive at this point to cite directly from Smith's (1973) summary of the major points in Bowman's (1971) work concerning the gamma motor system:

> "Most muscles, including many but not all of the muscles innervated by the cranial nerves, contain small fusiform receptors known as muscle spindles. The spindles are located mechanically in parallel with the muscle fibers, which are called extrafusal

fibers in this context. The spindle sends information back to the central nervous system (CNS) over two types of fibers: large diameter group Ia fibers, and somewhat smaller group II fibers. The afferent information, which can arise for example when the spindle is stretched, indicates in a relative sense both the length of the muscle (group Ia and group II fibers) and the rate of change of muscle length (group Ia fibers only). Extrafusal muscle fibers receive their motor innervation from large diameter alpha fibers, which are axons of alpha motoneurones in the spinal cord or brainstem motor nucleus. Spindles also contain muscle (contractile) fibers, called intrafusal fibers. They receive their innervation from gamma fibers (axons of gamma motoneurones). The gamma innervation of the spindle is a complex issue, and is not at present completely understood. The alpha and gamma motor systems in mammals are anatomically and functionally distinct. The spindle with its associated afferent and motor nerve supplies is considered to be one of the most important mechanisms regulating the stability and accuracy of muscle contraction. An accurate understanding of the functioning of this system is thus obviously important for detailed neuromuscular studies of the speech production apparatus." (Smith, 1973, p. 172)

Of similar import is the work of Merton (1953; cited by Abbs, 1973), who claims that the spindle motor system can produce output proportional to the length error between extrafusal and intrafusal fibre systems. The error signal is transmitted to the motoneurons of the extrafusal fibres as "negative feedback". The muscle length is then indirectly controlled by contraction of the spindle fibres. Abbs (1973) goes on to note that a common misunderstanding of the spindle motor system for feedback in speech is that many investigators feel it to be a peripheral neuromotor network. In support of this contention, Abbs cites Mortimer and Akert's (1961) findings from research with primates which confirms that "gamma motoneurons have discrete areas of cortical representation very similar to those of alpha motoneurons and in some cases the two types are excited by the same cortical region. Such representation suggests, not a diffuse facilitory action from the spindle motor system, but a detailed cortically controlled function" (Abbs, 1973, p. 178).

It is perhaps the role of the cerebellum, which receives afferent muscle impulses from the brainstem and cortex, to coordinate motor and sensory activity (cf. Ruch et al., 1967; cited in Abbs, 1973, p. 178). From pathological conditions of the cerebellum it has been noted to control precision of rate, range, force and direction of voluntary motion.

In sum:

> "the spindle motor system is not simply a peripheral control mechanism that serves only to modulate more central activities. The [available] data ... would suggest that this system has the functional representation to interact with alpha motoneuron systems in the cortical and subcortical generation of speech movements. Indeed, Granit, as early as 1955, suggested existence of separate but interacting control

of gamma and alpha motoneurons at higher neuromotor centers." (Abbs, 1973, p. 178)

Abbs mentioned the hesitation of many speech-production researchers to assign the role of "detection and correction" of speech errors to the gamma-loop because of the relatively long delay time involved. The contribution of primary spindle afferent fibres and the specific role of the spindle motor system in movement control are considered. As mentioned above, group Ia afferent spindles relay information on rate of length change of the muscle, and thus length can be anticipated and problems of overshoot or oscillation avoided. Consideration of the role of these group Ia fibres, especially that of feedback delay, leads to an awareness of a servofunction for the gamma efferent system (cf. Abbs, 1973, p. 179). In this regard Fairbanks (1954) postulated a model for speech in terms of a servosystem (see also Section 1.6 below); such a model compares output to input and adjusts input accordingly in this closed-loop system. Von Euler (1966; cited in Abbs, 1973, p. 179) supports the notion of a servosystem with work on muscle spindles in the intercostal muscles of the chest, in which firing rate of alpha motoneurons increases with increased rate of respiration. Thus,

> "A continuous error signal that modulates centrally generated alpha activity could provide a basis for the continuous correction of intended muscle length as set by the independently activated gamma motor fibers." (Abbs, 1973, p. 179)

In spite of its predictive powers, a simple servo model cannot explain continuous control in production of short rapid muscle movements, such as those necessary for speech. The delay "loop" in humans is around 20-80 ms (cf. Campbell, 1968; Alston et al., 1967; Sears & Newsome Davis, 1968: all cited in Abbs, 1973, p. 179), and some speech movements are completed in less time, such as those involving tongue tip, which are often initiated and completed in less than 50 ms.

Stark (1968; cited in Abbs, 1973, p. 180) has suggested that the spindle (gamma) system is used when "continuous" control is required, and the alpha system may operate when high speed or ballistic-type movement is required. Research by many investigators (for details, see Abbs, 1973, pp. 180-181) led Abbs to summarize the possible role of the spindle motor system as a "variable" servo model, in the following terms:

> "the spindle motor system operates (1) when the muscle is disturbed in an isometric state [i.e., when the ends of the muscle are fixed in place and increase in tension occurs without appreciable increases in length] or during controlled isotonic contraction [i.e., when the total muscle is of equal tension] by an unexpected force (that is, the spindle system attempts to maintain length or a certain rate of change in length), (2) to develop speed in the initiation of contraction, and (3) to provide antagonistic facilitation to damp movement and prevent overshoot." (Abbs, 1973, p. 181)

1.32 Auditory Feedback

The first suggestion that auditory feedback may be involved in speech monitoring and production, in terms of the effects of delayed auditory feedback on speech, was suggested by Lee (1950, 1951) and Black (1951). Other speech production models employing auditory feedback were proposed by Fairbanks (1954) and Chase (1958). Their models will be discussed below in Section 1.6.

Van Riper (1971) speculates about the role of auditory feedback for speech production. Acknowledging the controversy between continuous and intermittent monitoring of speech, Van Riper views speech as operating like a servosystem under ordinary conditions and claims that:

> "Information about the speech output is returned to the central integrating mechanism through six auditory channels, via the right and left feedback routes from (1) airborne side-tone, (2) bone-conducted side-tone, and (3) tissue connected side-tone. Other feedback signals come from the kinesthetic-tactile-proprioceptive sensors on both sides of the body. Stromsta (1962) showed that auditory feedback signals in these different channels arrive at markedly different times and that the temporal information-processing of speech output by the brain is very complex. Some central mechanisms for integrating all these feedback signals must be present, although their nature is not yet known." (Van Riper, 1971, p. 383)

Hardy (1970) discussed the importance of auditory monitoring for maintenance of speech production. He considered the congenitally deaf who do not usually develop intelligible speech and the adventitiously deaf (i.e., deafened after learning to speak) who show a slow deterioration of the speech process (cf. Sataloff, 1966; cited by Hardy, 1970, p. 50). Hardy supported the belief (after Chase, 1958, and others) that the auditory signal is part of the total sensory experience used in generating speech with complete information about speech musculature patterning contained in the speech "target".

Lombard (cited in Hardy, 1970, p. 51), who blocked auditory feedback by masking the subjects' speech with high intensity noise, found that -- other than raising the intensity level of the voice -- speakers showed little disruption of articulation. This led Hardy to reject Fairbanks's (1954) model of closed-loop feedback for speech, where the system would be dependent on auditory feedback at all times. Perhaps the speaker, in raising his voice, is adjusting his auditory feedback level so that it is audible under such conditions, at least via bone conduction. The adventitiously deafened speaker does not lose speech immediately, in spite of the auditory feedback having been lost. Hardy's conclusion may shed some light on these considerations:

> "it must be concluded that intraoral sensations can provide cues for positioning of the speech musculature once the appropriate patterning has been learned, and they can continue to do so in the absence of auditory feedback." (Hardy, 1970, p. 51)

1.33 Summary

To synthesize information from Section 1.2 on neurophysiology and from this section, a speculative comment from Konigsmark (1970), based on his knowledge of neural structures and their connections, is appropriate at this time:

> "The integrative activity resulting in speech probably begins in the cerebral cortex with a concept which can be vocalized. Broca's area in the cortex may then be influenced to initiate the speech process. Projections from this cortical region go to the motor cortex. From here a major projection courses to the motor nuclei involved in speech, that is, the hypoglossal nucleus, the nucleus ambiguus, the facial nucleus, and the motor nucleus of the Vth nerve. At the same time, fibers from the cerebral cortex project to the basal ganglia, and to the cerebellar cortex via the pontis [pons]. These projections probably function to smooth and to create the necessary motor tonus for vocalization. Projections from Broca's area to the respiratory motor area may coordinate this activity with speech.
>
> "Neurons in the hypoglossal nucleus, nucleus ambiguus, facial nucleus, and motor nucleus of the Vth nerve are played upon by projections from the precentral gyrus and by the cerebellar cortex via the red nucleus. Also, shorter connections interconnect these nuclei with one another, possibly aiding in their coordinated activity. Fibers from these nuclei act upon the musculature of the tongue, larynx, mouth, and jaw to produce speech.
>
> "Sensory endings in the mucosa and musculature of the tongue, larynx, mouth, and jaw are activated by touch, pressure, and position. This information

is fed into the dorsal horns of the spinal cord, the nucleus solitarius, and to the trigeminal sensory complex. These structures are also influenced by the cerebral cortex and reticular formation, possibly enhancing or dampening their activity, as the occasion demands. These sensory nuclei then project to the ventral posterior medial nucleus of the thalamus, and then to the postcentral gyrus of the cortex.

"Auditory feedback of what is being said projects to the transverse temporal gyri. From these gyri there are projections to the motor cortex, allowing a comparison of the results of speech and possibly influencing the motor production of speech." (Konigsmark, 1970, p. 17)

1.4 Speech Errors

As defined by Boomer and Laver (1968), a "slip of the tongue ... is an involuntary deviation in performance from the speaker's current phonological, grammatical or lexical intention" (Fromkin, 1971, p. 29).

Linguistically, speech errors have been studied for several reasons: (1) to provide important clues for language change, to provide a source for studying historical linguistic change, as suggested by Hermann Paul (Sturtevant, 1917; MacKay, 1970); (2) to understand better the speech production process via the mechanisms of speech (Boomer & Laver, 1968; Fromkin, 1968; Nooteboom, 1969); and (3) to draw a distinction between "competence" and "performance" and to demonstrate the reality of phonological units and rules (Fromkin, 1968).

Speech errors can occur whenever speech is used. Meringer (1908) recorded, along with the speech errors, the speaker's birthdate, his educational background, state of health, degree of tiredness, rate of speech, and the time of day at which such errors occurred, only to find that there was no correlation of any of these factors with the errors observed.

Several investigators have devoted time to classifying speech errors into such categories as misordering, omission or replacement of a unit (Boomer & Laver, 1968), or into phonemic vs non-phonemic errors (Nooteboom, 1969). Fromkin (1971) considers errors, not for purposes of classification, but as evidence for underlying units in speech, such as syllable, phoneme and feature.

From Fromkin's (1968) research on speech errors, one can see that such errors obey rules of the grammar and are not randomly generated. Her results show that: (1) features, segments and syllables make up units of production; (2) segments in a syllable are ordered, and this order is not violated in the production of an error; (3) morphemes or words of the same class (i.e., roots, affixes, etc.) are usually interchanged with one another; (4) intonation and primary word stress remain in the same position, regardless of the error; (5) morphological and phonetic or phonological constraints are placed on a word at different times in the generation of an utterance; (6) non-permissible phonetic sequences (i.e., those not characteristic of the language) do not occur; (7) semantic features may be displaced, resulting in a semantic error; and (8) words with similarity of phonological form are likely candidates for substitution as an error (cf. Fromkin, 1971).

As Fromkin (1971) and MacKay (1970) have noted, speech errors are more likely to occur between words that contain similar phonetic elements. It has also been noted that resultant errors of metathesis of two sequential phonemes in words (e.g., /æsk/ → /æks/) often seem to involve the sibilant /s/: "In a number of perception tests, the hiss (such as occurs with [s] is often 'misplaced'; i.e., it is difficult for subjects to judge where the noise occurs in an utterance" (Fromkin, 1971, p. 39). On the basis of such statements (and the data) one may suggest that words of similar phonological make-up involving the sibilant /s/ are the most likely to create speech errors.

Of the studies carried out on speech errors, those of Meringer and Mayer (1895) and Meringer (1908) are the most extensive in terms of number of errors and possible extra-linguistic correlations. Boomer and Laver (1968) collected over one hundred errors, and Fromkin (1971) reported over six hundred errors; but only Boomer and Laver tape-recorded their errors. Fromkin collected hers in an anecdotal fashion, generally with the speaker reporting after the fact what he had said and had meant to say. One major problem with Fromkin's method of data collection for a discussion at the molecular level of speech production is that subtle phonetic differences, or deviancies, in an error will (at least sometimes) be missed by the speaker and hence not reported, leading to false claims about the phonological/phonetic form -- and perhaps cause -- of a speech error.

In summary, speech errors have been used in various investigations to test diverse hypotheses, but no empirical examination of the error phenomena has itself been conducted. As a result of such considerations, it was felt that an examination of durations of segments in speech errors, specifically in word-initial consonant clusters -- an aspect of the problem not investigated by Meringer or others -- would prove informative and help elaborate hypotheses concerning the genesis of such errors.

1.5 Timing of Speech

Timing, for the purposes of this study, is defined as a sequential ordering of articulatory events, theoretically based on neuromuscular impulses sent by the brain to the articulators and programmed in some hypothetical unit, such as a phoneme, morpheme, or syllable.

It is not known from the literature whether timing of speech is regular; i.e., whether it remains constant, for example, from stress to stress within an utterance (cf. Ohala, 1970). Nor is it known when timing begins; i.e., whether a timing programme is set down when the first phoneme is uttered or as soon as the impulse is initiated in the brain. Ohala (1973) gives results which support the claims that there is no underlying speech rhythm or time programme, as claimed by Kozhevnikov and Chistovich (1965).

A number of investigators have considered timing in attempts to construct models of speech production, together with concomitant coarticulatory effects (e.g., Haggard, 1973; Kent & Moll, 1975; Lehiste, 1971; Ohala, 1970, 1973). These models can usually be classified in terms of systems which do or do not employ considerations of feedback in speech production (see also Section 1.6).

Timing is considered in this study only insofar as it is a possible determinant of the duration of segments which, it is tentatively hypothesized here, it constrains.

1.6 Models of Speech Production

1.60 Introduction

Research on speech errors, timing of speech and feedback has led to the formulation of numerous theories and models of speech production, several of which were briefly mentioned earlier. In general, models of speech production can be classified as either "closed-loop" or "open-loop": the former refers to a system which -- it is speculated -- employs feedback to regulate and adjust speech, the latter to a system which does not. The closed- *vs* open-loop distinction has been variously designated as a "chain" vs "comb" model (Bernstein, 1967), "sequential chain" vs "plan" model (Lenneberg, 1967), "hypothesis 1" vs "hypothesis 2", using efferent and afferent impulses (Kozhevnikov & Chistovich, 1965), and "peripheral feedback or chaining" vs "preprogramming" model (Kent & Moll, 1975). Our discussion begins with consideration of the theories which support an open-loop hypothesis.

1.61 Open-Loop Models

As noted above, an open-loop system specifies that commands are issued to the articulators at regular intervals to produce speech, but that feedback is not employed to regulate its production. One of the more innovative studies originated with Kozhevnikov and Chistovich (1965), who tested the rationale for using either of the two models to account for the sequential generation of syllables. Defining a "syntagma" as a sentence or phrase connected by meaning and articulation and pronounced on a single output, they varied the rate of speech and stress of the syntagma and found the syllable to be the unit which remained relatively constant under such variation. Their hypothesis was that if an open loop is in operation, the total variance over the time interval will be less than the sum of the variances of the component intervals, and their data support this claim.

MacNeilage (1970) initially supports an open-loop system, claiming that command patterns issued to the muscles would not wait for information from the articulator reaching a given target in order to control the following movement appropriately. However, he points out that an open-loop system would require storage of a vast amount of information on phoneme-to-phoneme transitions; and since this is not practical, he suggests that closed-loop control is more probable, based on neurophysiological research on the gamma motor system (cf. Section 1.3).

Fromkin's (1971) model of speech production shows a possible ordering of events in generation of an utterance and accounts for production of errors, as well as correct utterances, but it shows no relationship of these with any feedback mechanism. Since there is no mention of feedback, especially concerning error utterances in which the error is "caught" and then corrected, it can be assumed that Fromkin's model is more closely associated with an open-loop than with a closed-loop system.

1.62 <u>Closed-Loop Models</u>

The earliest notions of a closed-loop model, following cybernetic theory (i.e., a system which employs feedback by which to modify subsequent productions within a given utterance) are provided by Lee (1950) and Fairbanks (1954). Lee proposed a system of loops: articulation loops, monitoring phonemes via tactile and kinesthetic means, and voice loops, monitoring syllables via auditory feedback; both of these operate on volition and reflex systems. Fairbanks interpreted the speech system as a closed-loop servosystem, in which the output is compared with input via bone-conducted and air-conducted auditory feedback and which manipulates the production mechanism so that output will have the same functional form as input.

Chase's (1958) model of speech production likewise incorporates a servosystem and auditory feedback and evinces the same flaw as Fairbanks's model, namely that monitoring speech solely by means of auditory feedback would mean that speaking would be impossible (or at least inordinately difficult) in an extremely noisy environment, since feedback would be effectively masked. As many who work in industrial environments where there are exceptionally high noise levels will attest, this is not the case.

Kent and Moll (1975) posit a feedback model which assumes that timing of movements from higher centres depends on efferent and afferent signals received from a previous articulatory movement, in order to chain together speech segments. Their "preprogramming", or open-loop, model assumes that inherent timing control results in the timing of an articulatory movement's being affected by another articulatory movement which occurs either before or after it. While they initially interpret the data from their investigation of word-initial /spr-/ and /spl-/ clusters in terms of a closed-loop system, Kent and Moll note that responses from this feedback loop fail to reach consciousness and that articulation must therefore depend on "unconscious feedback-mediated responses" (1975, p. 319); it is not clear from their argument how feedback might ever be a conscious response. They reconsider their data in terms of an open-loop model, which they claim is the only way to account for: (a) variable duration of /s/ before /p/ in /sp-/ clusters, and (b) the ordering of /s/-release upon the gesture for /p/-closure, likewise in /sp-/ clusters. However, they do not say how an open-loop system would accomplish this: No resolution is forthcoming, except for the incidental non-comment that, whatever model one adopts, it will be "capable of fine and accurate control" (Kent & Moll, 1975, p. 321).

1.63 Summary and Discussion

Ohala (1970) tries to clear up some of the misconceptions surrounding the results and interpretation of Kozhevnikov and Chistovich's (1965) experiment, commenting that the methodology employed cannot reveal adequate information as to the presence or absence of feedback in the timing of speech. He proposes a reinterpretation of their closed- and open-loop hypotheses: The first is a system which is "'Timing Dominant' ... i.e., a system which maintains a tight time schedule perhaps at the expense of precise and thorough accomplishment of the gestures"; the second is one which is "'Articulation Dominant' ... i.e., a system which maintains precise and thorough performance of the gestures no matter how much time it takes" (Ohala, 1970, p. 143). He adds that both of these systems could either employ or not employ feedback in determining future articulatory events.

The major criticism leveled by Ohala (1970) against the methodology used by Kozhevnikov and Chistovich (1965), in which subjects were required to repeat the same utterance hundreds of times at different rates of speech, is that perhaps the subjects adopted a fixed rhythm and this could affect the underlying time schedule and generation of speech segments. Under such conditions an open-loop model, where timing commands are sent at fixed intervals, is more likely to be adopted as an explanation of their findings.

In terms of their data analysis, Ohala has criticized Kozhevnikov and Chistovich for their reliance on positive and negative covariances between phones in the determination of which model should be chosen; by covariance, here is meant that if an error is made in the duration of a phone, either positively or negatively, it is compensated for in the following phone, which finishes at the originally planned time, by being either shortened or lengthened, respectively. Ohala (1973) claims that variations in rate of speech of the test items, if consistently yielding positive covariances, would tend to support an open-loop model, which is indeed the one adopted by Kozhevnikov and Chistovich; if consistently negative covariances were obtained, on the other hand, a closed-loop model would suggest itself.

Ohala (1970) also questions the appropriateness of the closed-loop system for speech. He presents several arguments in favour of such a system and then proceeds to disprove all of them. He does, however, confirm the possibility of the use of short-term feedback to make quick adjustments in speech (by recourse to the results of his experiment on maximum jaw opening and velocity in the production of isolated words; cf. Ohala, 1970, pp. 122-141).

A model of speech production based on neurophysiological mechanisms which incorporates both open- and closed-loop systems has been proposed by Abbs (1973). He views the complete system as a "variable" servosystem (i.e., one which can employ feedback depending on the system's requirements). He claims that the gamma motor system (cf. Section 1.31) involved in feedback: (1) maintains the length or rate of change of length of a muscle, (2) is active in initiation of contraction, and (3) provides antagonistic actions to damp movement and prevent overshoot.

A speech perception/production model which is an active analysis process applied to the speech signal is the analysis-by-synthesis model proposed by Bell et al. (1961). The main part of this system is a generator capable of synthesizing all signals to be analysed. These signals are compared with signals to be analysed and an error measure computed. When a signal is synthesized that causes the error to reach the smallest value, this signal is stored. Components of the system are the filter set, spectrum generator, comparator, and strategy component. The designers claim that this system represents linguistic phenomena at various levels such as acoustic, phonological, morphological and syntactic. This system is mentioned here because it considers production, perception, and feedback of speech all in one efficient model.

On the assumption that feedback may or may not be involved in the production of speech (depending on various conditions, as yet unknown), arguments for one type of system or another are perhaps premature, given the lack of a proper empirical foundation. Much more research needs to be carried out in search of an answer to the problem, and attention should now be devoted to devising experiments which can adequately test for this intermittent feedback and to determining the role it plays in speech production.

CHAPTER 2

STATEMENT OF THE PROBLEM

There are few studies which examine models of speech production in conjunction with their possible neurophysiological bases, and even fewer in number are those investigations which incorporate considerations of feedback mechanisms into such models.

If we consider production of a speech error as a momentary breakdown, followed by a recalibration of the system enabling the correction of an error, and if we use such errors to hypothesize about certain aspects of speech production (including feedback and neurophysiological mechanisms), then perhaps it might be possible to provide further insight into the process of speech production.

Timing of speech is considered in relation to speech errors for the purpose of determining differences in duration of segments in repetition errors (of the form /s*C*/ ... /sC/), between /s*/ and /C*/ in an error production and /s/ and /C/ in a correctly produced repetition immediately following the error. Because the error is corrected immediately following its commission (although there may be a slight hesitation, to be described in Chapter 4), rate of speech is considered to remain constant and therefore not to affect significantly the length of consonants. Based on the foregoing considerations, comparisons are made between /s*/ and /s/ and between /C*/ and /C/.

The intent of the present study was to examine systematically specific aspects of speech errors through studying relationships of word-initial fricative plus stop consonant clusters, by:

(1) determining an efficient procedure for generating speech errors and for their subsequent analysis;

(2) obtaining a representative sampling of speech errors in word-initial fricative plus stop consonant clusters;

(3) providing a sample of speech at normal conversational rate containing no speech errors;

(4) classifying the types of speech errors found by consulting available descriptive accounts from previous research;

(5) evaluating the duration and timing relationships of certain of the word-initial clusters under investigation;

(6) considering the experimental results in terms of various models of speech production, feedback mechanisms, and speech perception models.

CHAPTER 3

METHOD

3.1 Pilot Studies

Prior to the main study, two pilot studies were conducted in order to ascertain a reasonably optimal approach to the collection and analysis of speech errors.

3.11 Pilot Study I

Subjects. Subjects for the first pilot study were four female university students, ranging in age from 21 to 25 years. All were native speakers of English with no demonstrable hearing or speech problems.

Stimulus Materials. Three types of stimulus materials were used in different combinations to determine which would produce the most fatigue of speech musculature and thus give the greatest number of speech errors:

(1) The first paragraph of "The Rainbow Passage" (cf. Fairbanks, 1960, p. 127; see Appendix A).

(2) Six occurrences of each of the clusters /sp-/, /st-/ and /sk-/ in initial position in words embedded in three separate paragraphs (see Appendix B).

(3) Twelve occurrences of each of the word-initial clusters specified in (2) above in words embedded in three sets of sentences, each referred to here as a "tongue-twister" (see Appendix C).

Procedure. Subject 1 was instructed to read the "Rainbow Passage" at normal speed, then to read the three paragraphs five times each consecutively as fast as possible. Subject 2 was instructed to read the "Rainbow Passage" at normal speed, and then to read the three "tongue-twisters" five times each as fast as she could. Subject 3 was told to read each of the paragraphs once at normal speed and then five times each as fast as she could. Finally, Subject 4 was required to read each of the "tongue-twisters" once at normal speed and then five times each as quickly as possible.

Results. Subject 1 produced three errors; Subject 2, thirteen errors; Subject 3, nine errors; and Subject 4, fifteen errors. Thus, more errors were produced by those subjects who had read the "tongue-twisters" than those who had not. In the light of these results, it was deemed necessary to conduct a second pilot study using only "tongue-twisters" as stimulus materials, in order to refine the experimental procedures.

3.12 Pilot Study II

Subjects. For this study subjects were two female university students, both 24 years of age, native speakers of English with no demonstrable hearing or speech difficulties.

Stimulus Materials. The stimulus materials consisted of the three "tongue-twisters" described in Section 2.11 above and given in Appendix C. Each sentence contained twelve occurrences of one of the word-initial clusters /sp-/, /st-/ or /sk-/.

Procedure. Each subject was required to read each sentence fifteen times as fast as possible. Only the last five repetitions were examined for speech errors, the first ten productions being considered as the "fatiguing" portion of the experiment.

Results. The number of errors produced by subjects reading "tongue-twisters" (in both pilot studies) was greater by a factor of two to one than those obtained through the use of any other passage. As such, they were considered to produce better "fatiguing" effects and were therefore chosen as stimulus materials for the experiment proper.

3.13 Discussion

The types of errors produced by the subjects in the pilot studies could be classified as follows: (a) syllable repetition, (b) phonetic substitution, (c) cluster repetition, (d) phoneme repetition, and (e) phoneme prolongation. These errors will be discussed more fully in Chapter 4.

It was felt that reading errors might have contributed to the percentage of the errors produced, but there appeared to be no objective way to extract speech errors under such conditions. In order to avoid the possibility of reading errors, given that the purpose of the investigation was to examine spontaneous productions, it was decided to have the subjects in the main study memorize the "tongue-twisters" rather than read them. Such a method has been described and employed by Kozhevnikov and Chistovich (1965).

3.2 Main Study

Subjects. Selected for inclusion in the study were twelve university students, ranging from 22 to 33 years of age, all of whom were native speakers of English. The hearing of all subjects was judged to be within normal limits, and none had deviant articulation patterns or anomalies of facial musulature.

Equipment. The equipment employed in this study included:
(a) For recording: a two-channel power supply (Brüel & Kjaer, Type 2803), associated with a Brüel & Kjaer one-inch microphone, Model 4145; and a Scully Model 280-2 tape recorder/reproducer. Recordings were made on Ampex 611 audiotape at 7.5 ips.
(b) For analysis: a Kay Sona-Graph, Model 7029A, a Siemens Oscillomink, and an Ampex Micro 50 cassette recorder/reproducer, utilizing studio-quality magnetic tape.

Stimulus Materials and Procedure. The stimulus materials employed included the three "tongue-twisters" described above and given in Appendix C.

Each subject was instructed to memorize one sentence at a time. When the subject indicated that she knew the passage well enough to recite it aloud without error or prompting, taping began, during which time a 12-inch mouth-to-microphone distance was maintained. Instructions to all subjects were as follows:

> "Repeat the sentence once at normal conversational rate in a normal conversational voice, and then fifteen times as fast as you can."

The subjects followed this procedure for all three "tongue-twisters", whereby the order of presentation of the three sentences was varied randomly for all subjects.

3.3 Analysis of Data

Tapes and Transcription. Tapes were transcribed for all twelve subjects by the experimenter using a modified version of the International Phonetic Alphabet (IPA). This analysis included the first production at normal speed and the fifteen test repetitions, including any errors, together with their immediate phonetic contexts. Tapes could be played back either at normal speed or at half-normal speed, in order to facilitate transcription. Errors were then coded (see Appendix D for examples of transcription and coding); and where necessary, spectrograms were produced to determine more clearly phonetic substitutions, epenthetic phones, and reversals within clusters.

Editing. Because of the extent of the data available, only data from the six subjects who made the most errors were edited and analyzed. Phonetic transcriptions were varified by having a trained phonetician transcribe errors from these six subjects. Editing was carried out via the Scully 280-2 tape recorder/ reproducer in conjunction with the Ampex Micro 50 cassette recorder, the experimenter isolating the normal-conversational-rate repetition of the "tongue-twisters", all errors, and their immediate contexts.

Oscillograms. Edited data were displayed visually on oscillograms, using the Siemens Oscillomink. This instrument displays a speech signal trace, duplex oscillogram trace, and a trace of the log of average speech power. Within the experimental arrangement is a Revox Model A77 tape recorder/reproducer, duplex oscillograph, Frøkjaer-Jensen Trans-Pitchmeter, and an intensity or speech power circuit. Oscillograms were produced at ten cm/s.

Insert Figure 1 about here.

Segmentation. Because of the rapidity of the subjects' speech, segmentation of the oscillograms proved somewhat difficult. Gross segmentation was carried out first. This was accomplished by marking off 10-cm sections on the oscillogram. The utterance on the tape was then timed with a stopwatch to a five-second mark, and at this point the oscillogram was matched by counting working backwards from every five to three to one second marks. When a subject's utterances were segmented at the gross level, a finer segmentation procedure was conducted.

One objective of the finer segmentation was to establish the time in milliseconds of the /s/-segments plus stop consonant

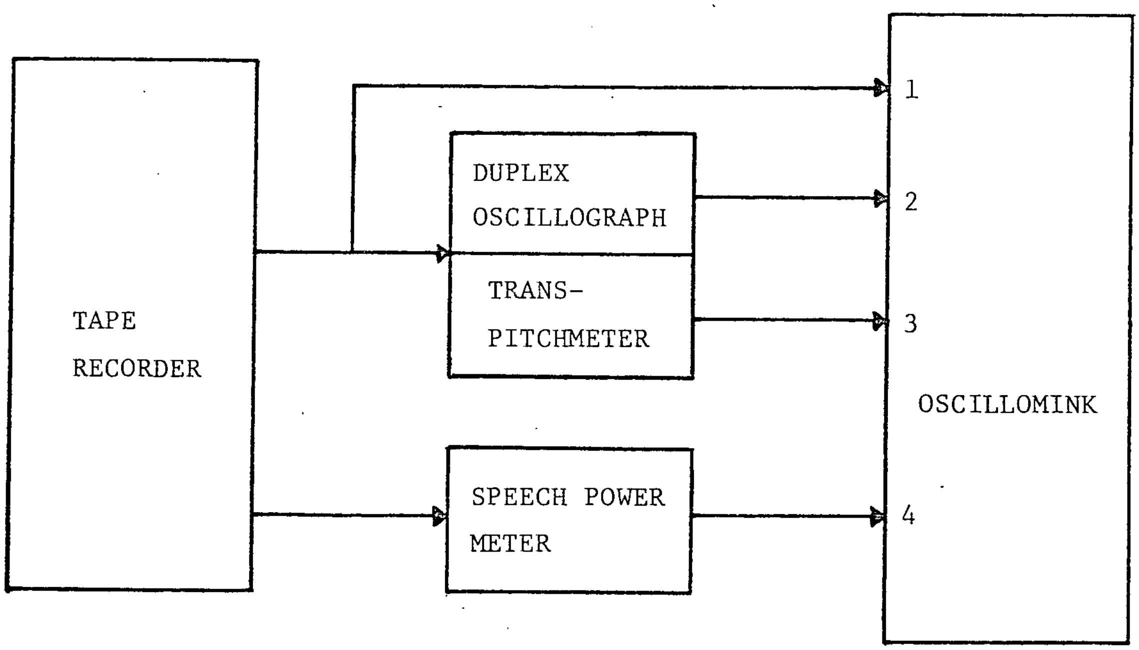

FIGURE 1. Block diagram of instrumentation used to produce oscillograms.

in: (a) the error segment, and (b) the correct productions of the utterance. Selection of /s/ + /C/ (stop) avoided many of the difficulties associated with segmentation of glides, resonants, vowels and phonemes within words which were excessively shortened, distorted or omitted because of the rapidity of the subjects' speech (cf. Lehiste & Peterson, 1959, for details).

Initial /s/ plus stop consonant plus vowel configurations can be segmented without difficulty by examination of the speech wave trace and the duplex oscillogram trace. The first trace, the speech signal, can be examined in conjunction with the negative amplitude of the duplex, or second, trace to determine an /s/; (cf. Lehiste and Peterson, 1959). The consonants /p/, /t/ and /k/ immediately following the /s/ can also be easily distinguished, since the trace of the speech signal and duplex both follow the zero-line. Vegetative sounds made by the subjects, such as snorts, breaths, lip smacks or clearing of the throat, made segmentation somewhat easier by adding natural pauses between often indistinct utterances.

The maximum error measurement which resulted from the above procedure was ten milliseconds, when one subject's data were measured again by the experimenter. This maximum error was present on only one of fifty measurements made.

Utterances which presented problems in segmentation (e.g., very rapid speech, or instances of dubious phonetic transcriptions) were, as noted above, classified by examining wide-band spectrograms.

CHAPTER 4

RESULTS

4.0 Introduction

Using the "tongue-twisters", each of which contained twelve instances of either /sp-/, /st-/ or /sk-/ in word-initial position (as described in Chapter 3 and given in Appendix C), a sample of the six subjects' speech under normal conversational conditions was obtained. Using the same "tongue-twisters" produced at a subjectively faster rate of speech was a procedure determined to be effective for generating speech errors in the pilot studies. A corpus of speech errors involving the word-initial clusters was also gathered and subsequently classified.

The six subjects produced a total of 228 errors, which could be categorized into six types, as follows (see also Table 1):

(1) Omission: deletion of a phoneme or syllable; e.g., spotted → -potted.

(2) Addition: epenthesis of a vowel or consonant; e.g., skimpy → kskimpy.

(3) Substitution: phonetic replacement of deviation from the target phoneme; e.g., skin → skim.

(4) Checked Hesitation: insertion of glottal stop as a pausal phenomenon; e.g., scallions → scallʔions.

(5) Prolonged Hesitation: unusual lengthening or prolongation of a phoneme; e.g., stalwart → st:alwart.

(6) Repetition: reproduction of a word, when the first production is halted following the first phoneme, cluster, syllable or entire word; e.g., spirited → spspirited.

Insert Tables 1 and 2 about here.

More than fifty percent of all errors obtained were of type (6), i.e., Repetition Errors. These in turn could be classified into six sub-types, when grouped according to the repeated segment or segments, as follows (see Table 2):

(1) Phoneme Repetition: production of the initial /s/, followed by a pause and then production of the full word; i.e., /s* (pause) sC.../.

(2) Cluster Repetition with a Pause: production with a pause between the initial cluster and repetition of the entire word; i.e., /s*C* (pause) sC.../.

(3) Cluster Repetition without a Pause: same as (2), but without a pause between the error cluster and the repetition; i.e., /s*C* sC.../.

(4) /sCV/-Syllable Repetition: production of an open syllable, followed by a pause and the production of the whole word; i.e., /s*C*V (pause) sC.../.

TABLE 1. Types and Numbers of Speech Errors Produced by Each Subject.

Subject	Error Type Omis-sion	Addi-tion	Substi-tution	Checked Hesita-tion	Prolong Hesita-tion	Repeti-tion	TOTAL
1	4	2	1	2	5	18	32
2	8	2	2	1	11	28	52
3	---	1	4	1	3	15	24
4	1	---	---	6	19	33	59
5	5	1	2	1	17	11	37
6	7	---	1	2	1	13	24
TOTAL	25	6	10	13	56	118	228

TABLE 2. Types and Numbers of Repetition Errors Produced by Each Subject.

Subject	Repetition Error Type Phoneme	Cluster w/ Pause	Cluster No Pause	Syllable sCV-	Syllable sCVC-	Entire Word	TOTAL
1	2	4	3	5	4	---	18
2	9	3	2	5	8	1	28
3	5	2	2	2	1	3	15
4	3	10	4	8	5	3	33
5	---	---	1	5	3	2	11
6	---	7	1	4	1	---	13
TOTAL	19	26	13	29	22	9	118

(5) /sCVC/-Syllable Repetition: production of a closed syllable, followed by a pause and the production of the entire word; i.e., /s*C*VC (pause) sC.../.

(6) Word Repetition: production of the entire word followed by its repetition.

The above repetition error types were examined with respect to the duration of the fricative plus stop consonant clusters, as well as to that of the individual segments of which they were composed.

4.1 Control Group Data

A speech sample at each subject's normal rate of speech was obtained; and the durations of the /sC-/ clusters, including internal segments, were then measured and analyzed. The duration data thus obtained provide norms for this study and and hereafter referred to as the Control Group Data (CGD). The summary statistics for these data are presented in Table 3.

Insert Table 3 about here.

The CGD were subjected to Bartlett's test for homogeneity of variance across subjects: The chi-square values obtained did not exceed the critical value for significance ($p > .01$); it was therefore assumed that the subjects' individual data could be pooled for purposes of further evaluation and analysis.

TABLE 3. Control Group Data (Normal Conversational Rate) for Word-Initial Consonant Clusters /sp-/, /st-/, /sk-/: Means and Standard Deviations of Segmental Durations (in ms).

Segment		Subject 1 N=12	2 N=12	3 N=12	4 N=12	5 N=12	6 N=12	Mean N=72
$/s_p/$	$\bar{X}$ =	76.67	114.17*	140.00	89.17	91.67*	74.17	97.64
	sd=	17.23	53.17	32.75	31.47	31.86	21.09	39.56
/p/	$\bar{X}$ =	80.83	80.00*	90.83	82.50	91.67*	85.00	85.14
	sd=	12.22	11.28	18.81	23.01	9.37	16.79	16.08
/sp/	$\bar{X}$ =	157.50	194.17*	230.83	171.67	183.33*	159.17	182.78
	sd=	20.39	57.91	46.41	49.70	34.20	23.14	46.90
$/s_t/$	$\bar{X}$ =	100.83	117.50	123.33	87.50	80.00	93.33	100.42
	sd=	15.64	21.79	28.39	28.32	20.45	27.08	28.01
/t/	$\bar{X}$ =	40.83	39.17	43.33	57.50	44.17	45.00	45.00
	sd=	18.32	10.84	11.55	8.66	13.79	17.84	14.73
/st/	$\bar{X}$ =	141.67	156.67	166.67	145.00	124.17	138.33	145.42
	sd=	29.80	21.03	35.51	29.70	24.29	24.43	30.11
$/s_k/$	$\bar{X}$ =	93.33	107.50	125.00	85.00*	95.83	92.50	99.86
	sd=	24.25	26.33	25.05	21.95	17.30	18.65	25.37
/k/	$\bar{X}$ =	57.50	63.33	61.67	61.67*	59.17	56.67	60.00
	sd=	20.51	11.55	11.93	12.67	7.93	13.03	13.22
/sk/	$\bar{X}$ =	150.83	170.83	186.67	146.67*	155.00	149.17	159.86
	sd=	43.16	29.38	28.39	28.71	21.53	29.68	33.04

* Due to subject error, N=11 for these entries; mean value added in each case to yield N=12 in order to standardize observations across all subjects.

Several analyses of variance were carried out: A two-way classification demonstrated significant differences among subjects, segments and the interaction of these (see Table 4); but with respect to subject-by-cluster interaction, no such significance could be found (see Table 5). Under each of these analyses, a Newman-Keuls test (with $p < .05$) was carried out, which indicated that the individual clusters comprised homogeneous subsets, as did each of the stop consonants, whereby all initial /s/-segments fell into the same subset (see Tables 4a and 5a).

Insert Tables 4 - 7a about here.

One-way analyses of variance were also carried out with the data from all subjects pooled. As expected, the segments and clusters showed significant differences, and the Newman-Keuls test (with $p < .05$) demonstrated the same homogeneous subset groupings as were found in the two-way classification (see Tables 6 and 6a, 7 and 7a).

A general interpretation of the CGD can be made by consideration of the overall means (as given in the last column of Table 3), which are presented graphically in Figure 2. The duration of /s/ before any of the stop consonants was approximately 100 ms, with /s/ before /p/ being slightly shorter. The mean duration of the stop consonants ranges from 45-85 ms, with /t/ having the shortest and /p/ having

TABLE 4. Summary of Analysis of Variance: Control Group Data -- Cluster Segments and Subjects.

Source of Variation	Sum of Squares	d.f.	Mean Square	*F*	*p*
Subjects (S)	29585.18	5	5917.04	12.566	<.0001
Segments (P)	198937.95	5	39787.59	84.497	<.0001
S x P	42581.48	25	1703.26	3.617	<.0001
Error	186466.65	396	470.88		

TABLE 4a. Newman-Keuls Summary Table: Control Group Data ($p < .05$) -- Cluster Segments and Subjects.

Homogeneous Subsets: Subjects		Homogeneous Subsets: Cluster Segments	
(1)	6, 1, 5, 4	(1)	$/s_p/$, $/s_k/$, $/s_t/$
(2)	2	(2)	/t/
(3)	3	(3)	/k/
		(4)	/p/

TABLE 5. Summary of Analysis of Variance: Control Group Data -- Clusters and Subjects.

Source of Variation	Sum of Squares	d.f.	Mean Square	*F*	*p*
Subjects (S)	59170.01	5	11834.07	10.412	< .0001
Clusters (C)	51112.03	2	25556.02	22.484	< .0001
S x C	13860.18	10	1386.02	1.219	> .25
Error	225049.98	198	1136.62		

TABLE 5a. Newman-Keuls Summary Table: Control Group Data ($p < .05$) -- Clusters and Subjects.

Homogeneous Subsets			
Subjects		Clusters	
(1)	6, 1, 5, 4	(1)	/st/
(2)	2	(2)	/sk/
(3)	3	(3)	/sp/

TABLE 6. Summary of Analysis of Variance: Control Group Data -- Cluster Segments, One-Way Classification.

Source of Variation	Sum of Squares	d.f.	Mean Square	*F*	*p*
Segments	198937.95	5	39787.59	65.535	$< .0001$
Error	258633.31	426	607.12		

TABLE 6a. Newman-Keuls Summary Table: Control Group Data ($p < .05$) -- Cluster Segments.

Homogeneous Subsets: Segments	
(1)	$/s_p/$, $/s_k/$, $/s_t/$
(2)	/t/
(3)	/k/
(4)	/p/

TABLE 7. Summary of Analysis of Variance: Control Group Data -- Clusters, One-Way Classification.

Source of Variation	Sum of Squares	d.f.	Mean Square	F	p
Clusters	51112.03	2	25556.02	18.26	< .0001
Error	298080.53	213	1399.44		

TABLE 7a. Newman-Keuls Summary Table: Control Group Data ($p < .05$) -- Clusters.

Homogeneous Subsets: Clusters
(1) /st/
(2) /sk/
(3) /sp/

the longest duration. The clusters range in mean duration from about 145 ms for /st/ to 183 ms for /sp/, with /sk/ in between at 160 ms. The duration of the clusters is consistently of the same order as with the individual stop segments, due to the similarity of the durations of the /s/-segments.

Insert Figure 2 about here.

4.2 Experimental Group Data

The six Repetition Error classes (hereafter called Experimental Groups #1 - #6) were examined separately by subject for segments and clusters, the summary statistics for which are presented in Table 8.

Insert Table 8 about here.

In Experimental Group #1 (Phoneme Repetition), /s/ is consistently shorter the /s*/ (i.e., the error production), regardless of the stop consonant which follows. The segment /p/ is the longest stop consonant (ca. 100 ms), followed by /k/ (87.5 ms) and finally by /t/ (50 ms); this is consistent with the findings for the CGD. The relative ordering of the cluster durations finds /sk/ to have the longest mean duration and /st/ the shortest; this differs from the CGD in that

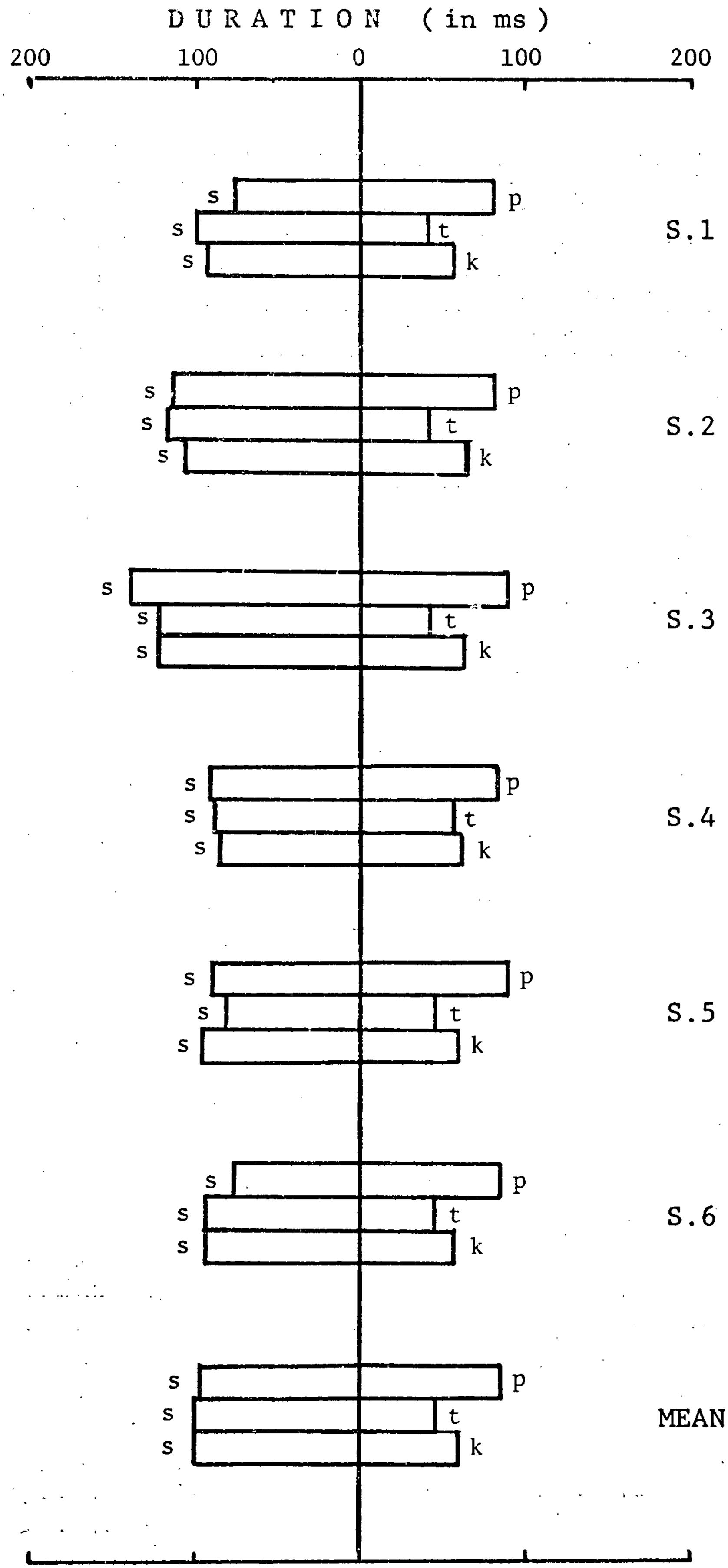

FIGURE 2. Control Group Data: Segment and Cluster Duration. Mean values for each subject.

TABLE 8. Summary Statistics for Segmental and Cluster Durations (in ms), in Six Repetition Error Groups (termed Experimental Groups). (Mean / Standard Deviation / Number of Observations)

Experimental Group		s*p* , sp	s*t* , st	s*k* , sk
#1: Phoneme Repetition (s* - sC)	s*:	136.67/ 66.58/ 3	127.50/ 92.15/ 4	162.50/154.17/12
	s:	90.00/ 10.00	107.50/ 57.37	154.17/122.14
	C*:	-----	-----	-----
	C:	100.00/ 50.00	50.00/ 24.49	87.50/ 61.52
	s*C*:	-----	-----	-----
	sC:	190.00/ 45.83	157.50/ 71.36	241.67/136.90
#2: Cluster Repetition (with Pause)	s*:	120.00/ 62.05/ 5	128.00/ 35.64/ 5	118.75/ 45.73/16
	s:	68.00/ 16.43	100.00/ 40.00	116.25/ 31.60
	C*:	208.00/110.77	318.00/261.29	229.37/227.29
	C:	82.00/ 4.47	52.00/ 19.24	66.87/ 29.15
	s*C*:	328.00/129.31	446.00/273.28	348.12/242.63
	sC:	150.00/ 15.81	152.00/ 48.68	183.12/ 48.82
#3: Cluster Repetition (No Pause)	s*:	128.00/ 48.68/ 5	140.00/ 56.57/ 2	158.33/105.53/ 6
	s:	118.00/ 65.35	95.00/ 21.21	131.67/ 36.01
	C*:	462.00/308.50	270.00/296.98	376.67/164.03
	C:	45.00/ 25.50	40.00/ 28.28	65.00/ 10.49
	s*C*:	590.00/343.29	410.00/240.42	535.00/231.58
	sC:	163.00/ 85.41	135.00/ 7.07	196.67/ 41.79
#4: sCV-Syllable Repetition	s*:	118.33/ 51.15/ 6	87.69/ 33.20/13	103.00/ 32.68/10
	s:	116.67/ 54.65	90.77/ 33.78	102.00/ 39.38
	C*:	160.83/154.93	79.23/ 57.22	92.00/ 41.04
	C:	70.00/ 17.89	43.08/ 20.97	60.00/ 23.09
	s*C*:	279.17/156.67	166.92/ 81.69	195.00/ 54.42
	sC:	186.67/ 61.21	133.85/ 50.42	162.00/ 53.71
#5: sCVC-Syllable Repetition	s*:	116.67/ 46.19/ 3	80.00/ 19.27/ 8	121.82/ 47.08/11
	s:	120.00/ 50.00	105.00/ 51.27	105.45/ 26.22
	C*:	123.33/ 77.67	40.00/ 13.09	148.18/254.67
	C:	90.00/ 26.46	41.25/ 16.42	45.45/ 13.68
	s*C*:	240.00/122.88	120.00/ 22.04	270.00/286.23
	sC:	210.00/ 72.11	146.25/ 55.79	150.91/ 35.34
#6: Word Repetition	s*:	115.00/ 49.50/ 2	60.00/ 0.00/ 2	108.00/ 32.71/ 5
	s:	75.00/ 21.21	110.00/ 14.14	108.00/ 22.80
	C*:	115.00/ 7.07	55.00/ 7.07	82.00/ 17.89
	C:	80.00/ 28.28	80.00/ 28.28	70.00/ 10.00
	s*C*:	230.00/ 56.57	115.00/ 7.07	190.00/ 47.43
	sC:	155.00/ 49.50	190.00/ 14.14	178.00/ 30.33

/sk/ and /sp/ are reversed in order with respect to duration. Because of the phonetic shape of this experimental group, no segment /C*/ or cluster /s*C*/ exists (see also Figure 4).

In Experimental Group #2 (Cluster Repetition with a Pause), the error productions /s*/, /C*/ and /s*C*/ are all longer in duration than the corrected repetitions /s/, /C/ and /sC/, respectively. Considering the error clusters: /s*t*/ is the longest (446 ms) and /s*p*/ the shortest (328 ms), a finding which is not consistent with the other experimental groups. The cluster /sk/ has the longest duration (183 ms) and /sp/ the shortest (150 ms), which is not consistent with the CGL (see Figure 3).

Experimental Group #3 (Cluster Repetition without a Pause) patterns after the above group, where /s*/, /C*/ and /s*C*/ are consistently longer in duration than /s/, /C/ and /sC/, respectively. In this group, /C*/ is definitively longer than in any other experimental group: the cluster /s*p*/ is the longest (590 ms) and /s*t*/ the shortest (410 ms), while /sk/ is the longest (197 ms) and /st/ the shortest (135 ms) of the corrected productions. The cluster /st/ is the only one which is consistent with the CGD (see also Figure 3).

Experimental Group #4 (/sCV/-Syllable Repetition) also shows all error segments to be longer than non-error segments. Here, /s*p*/ and /sp/ are the longest and /s*t*/ and /st/ the shortest clusters.

In Experimental Group #5 (/sCVC/-Syllable Repetition), the non-error segments and clusters are longer than the error segments and clusters, respectively. This differs from all other groups discussed thus far. Of the non-error productions, /s_p/, /p/ and /sp/ have the longest durations, while /s^*_k/, /k*/ and /s*k*/ show the longest durations of the error segments and clusters (see Figure 3).

Experimental Group #6 (Word Repetition) demonstrates /s_t/, /t/ and /st/ to have the longest durations, while /s_p/, /p/ and /sp/ have the shortest. Of the error data, /s^*_p/, /p*/ and /s*p*/ are the longest, while /s^*_t/, /t*/ and /s*t*/ are the shortest, thus conflicting with the results for the CGD (see also Figure 4).

Insert Figures 3 and 4 about here.

Insert Table 9 about here.

In order to generalize observations, the data from Experimental Groups #2-#5 were pooled, the results of which are presented in Table 9. This summary illustrates that /s*p*/ at 368 ms and /sp/ at 174 ms exhibit the longest cluster durations, while /s*t*/ at 221 ms and /st/ at 141 ms

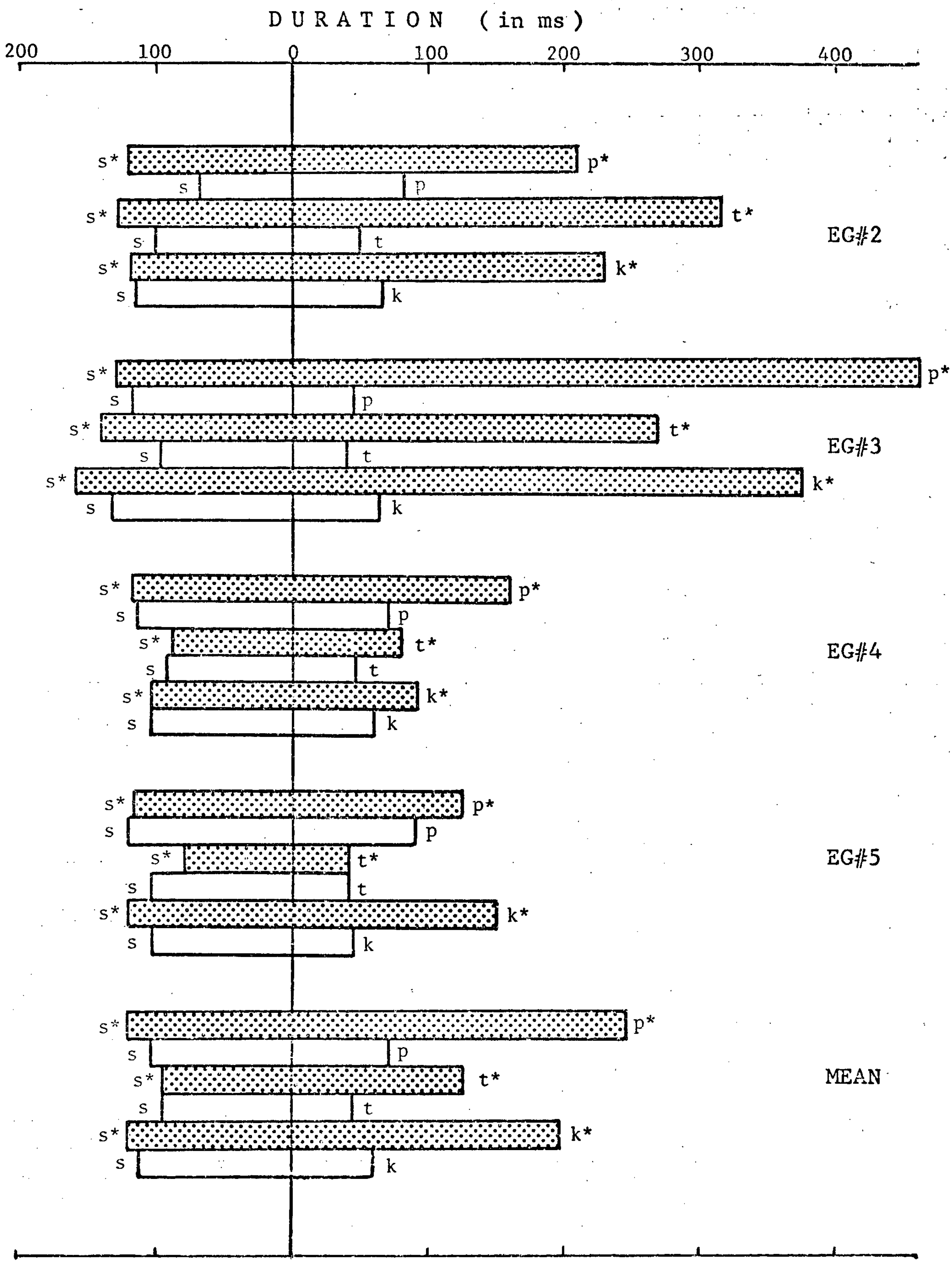

FIGURE 3. Experimental Groups #2-#5: Segment and Cluster Duration. Mean values for error and correct productions.

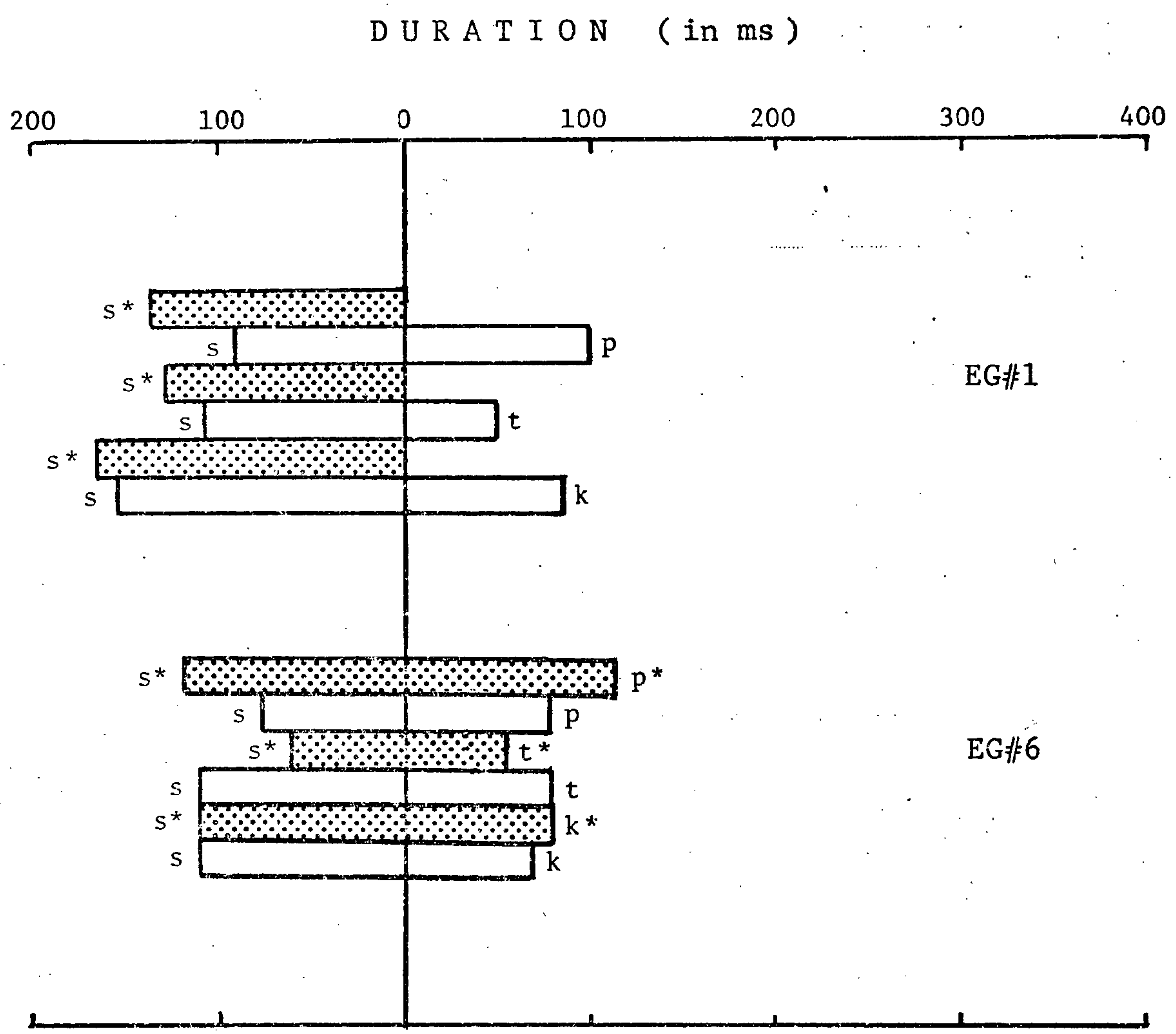

FIGURE 4. Experimental Groups #1 and #6: Segment and Cluster Duration. Mean values for error and correct productions.

TABLE 9. Summary of Repetition Error Data for Experimental Groups #2-#5: Segmental and Cluster Durations (in ms). (Mean / S.D. / N)

Segment	Error Production	Segment	Correct Production
$/s^{*}_{p}/$	121.05 / 48.64 / 19	$/s_{p}/$	104.74 / 51.25 / 19
$/s^{*}_{t}/$	96.43 / 36.54 / 28	$/s_{t}/$	96.79 / 38.50 / 28
$/s^{*}_{k}/$	121.40 / 55.75 / 43	$/s_{k}/$	112.33 / 33.23 / 43
/p*/	246.58 / 222.49 / 19	/p/	69.74 / 24.41 / 19
/t*/	124.29 / 162.74 / 28	/t/	43.93 / 19.12 / 28
/k*/	197.21 / 213.81 / 43	/k/	59.53 / 23.50 / 43
/s*p*/	367.63 / 240.61 / 19	/sp/	174.47 / 61.30 / 19
/s*t*/	220.71 / 180.96 / 28	/st/	140.71 / 48.45 / 28
/s*k*/	318.60 / 242.54 / 43	/sk/	171.86 / 47.42 / 43

exhibit the shortest; this is consistent with the CGD findings. The segments $/s^*_k/$ at 121 ms and $/s_k/$ at 112 ms are only slightly longer in the latter case than $/s^*_p/$ at 121 ms and $/s_p/$ at 105 ms, with $/s^*_t/$ at 96 ms and $/s_t/$ at 97 ms remaining the shortest. Also consistent with the CGD is the finding that /p*/ at 247 ms and /p/ at 70 ms are the longest stop segments, whereby /t*/ at 124 ms and /t/ at 44 ms are the shortest.

The following general observations may be made:

(1) Segments and clusters in the error, or first, productions are longer than their respective counterparts in the corrected, or second, productions.

(2) The shortest clusters are /s*t*/ and /st/, which also contain the shortest segments $/s^*_t/$, /t*/, $/s_t/$ and /t/.

(3) The longest /s/-segments are $/s^*_k/$ and $/s_k/$.

(4) The longest stop consonants are /p*/ and /p/.

(5) The longest clusters are /s*p*/ and /sp/.

If we consider the results of overall means obtained in the correct productions in light of the CGD findings, we note that:

(1) The duration of /s/ before any stop consonant was approximately 100 ms in the CGD and 104 ms in the Experimental Group Data (EGD).

(2) The duration of the stop consonant was 63 ms in the CGD and 58 ms in the EGD.

(3) The duration of the cluster was 163 ms for both CGD and EGD.

The ranking of the segments and clusters within their respective groups also remains about the same. An interesting observation might be made regarding the similarity of these means: Under control conditions, subjects were speaking at a normal rate, while the same subjects, when speaking under experimental conditions, were speaking as fast as they could. This should, it would seem, make the durations of the EGD segments and clusters shorter than those of the CGD; but such was not the case. It is reasonable to suppose from these results that once an error in a first production is made, perhaps the rate returns to the normal conversational rate until the subject can once again pick up speed.

Summary statistics were also derived for the pause, or delay time, between error and correct productions. These statistics show that the pause can be eliminated, but that a lengthening of the consonant before the pause, when it occurs, can have a mean value as great as 967 ms (such as in Phoneme Repetition). Disregarding the cases where no pause occurs, the pause can be as short as 20 ms or as long as 6550 ms (with little or no lengthening of the segment before it).

CHAPTER 5

DISCUSSION

5.0 General Considerations

The intent of the present study was to examine systematically specific aspects of speech errors through studying relationships of word initial fricative plus stop consonant clusters. The categories into which error utterances are classified in the literature pertaining to speech errors (specifically Boomer & Laver, 1968) were not adequate to describe errors produced by subjects in this study. In addition to misordering, omission and replacement of segments (cf. Boomer & Laver, 1968), addition, hesitation and repetition errors were produced. These latter categories were found by consulting the literature on delayed auditory feedback (DAF) and were adapted from the categories set up by Fairbanks and Guttman (1958), since they proved to be the most applicable to the present study.

Approximately fifty percent of all errors produced by subjects in this study were of the repetition type. Categorization of these errors was based on phonetic transcriptions carried out by the experimenter and a trained phonetician and supplemented by spectrographic analysis where necessary (for a list of the errors and the phonetic transcriptions, see Appendix D). This yielded six categories of repetition errors based on their phonetic forms.

Speculation as to why repetition errors are corrected and errors in the other categories are not led to examination of the first productions for phonetic deviancies *vis-à-vis* the subsequent correction productions. It was found that only about one third of the repetition errors could be considered as corrections because of phonetic abnormalities in the first, or error, production. The phenomenon of excessive length associated with the initial cluster and its component segments was often noted, but it was not considered as a phonetic abnormality. It was observed that the first phoneme or cluster in the error syllable or word was somewhat longer in duration than might have been subjectively expected. This observation led in turn to comparison of the durations of the initial clusters and their segments in the error production with the immediately following cluster production, or correction. In order to measure objectively and compare these durations, oscillograms were produced and measured for each subject's normal production of the "tongue-twisters", then for each subject's errors which occurred in subsequent rapid productions. The results obtained were subjected to statistical analysis (as detailed in the previous chapter). Findings specific to the control group data will be discussed first.

5.1 Discussion of the Control Group

The analysis of variance carried out on the control group data showed significant differences between subjects and between phonemes, as well as clusters (cf. Tables 4-7). The

Newman-Keuls test grouped four of the subjects into one homogeneous subset, while Subject 2 and Subject 3 were each separately grouped. The segregation of these two subjects may have resulted from the fact that both spoke with subjectively more precise articulation and somewhat more slowly than the other subjects (but also differently enough from one another to be grouped separately). Such factors would tend to lengthen segments and clusters in the speech of these two subjects and thus account for the differences in the Newman-Keuls test (cf. Tables 4a and 5a).

The four significant subsets for segments exhibited by the Newman-Keuls test (cf. Tables 4a and 6a) and the three subsets for clusters (cf. Tables 5a and 7a) segregate (a) all /s/-segments as a group from /p/, /t/ and /k/, each of which are also grouped separately, and (b) the clusters, /st/ from /sk/ from /sp/. Examination of the means (cf. Table 3) indicates that /s/ before /p/ is slightly shorter than /s/ before /t/ or /k/, with /t/ being the shortest stop and /p/ the longest, consistent with the data reported by Schwartz (1970). Based on the means and the Newman-Keuls test groupings in this study, it is reasonable to speculate that, since duration of the sibilant is similar in each context, it is the stop consonant which ultimately determines the duration of the cluster as a whole. Indeed, when one examines the means and the subset orderings for the stop consonants and for the clusters, the same relationship holds; i.e., /p/ is longer than /k/, which is longer than /t/, and /sp/ is longer than /sk/, which is longer than /st/.

The above findings are in agreement with those of Borden and Gay (1975) with respect to the groupings and relative orderings of segments comprising word-initial clusters; due perhaps to their subjects' producing words in isolation, their values for the durations of these segments are somewhat larger in all cases. Inconsistent with their findings is the relative ordering of cluster lengths, since their data indicate that /sk/ is slightly longer than /st/, which is longer than /sp/. This suggests that for their three subjects the sibilant may be a greater determiner of cluster length, which is in contrast with the results of the present study.

5.2 Discussion of the Experimental Groups

In experimental groups #1-#6 the duration of the first, or error, production was contrasted with the second, or corrected, production. Because of the differences in rate of speech and other uncontrolled variables, these speech error data cannot be legitimately compared with the control group findings; however, the general trends in the two groups of data can be compared in order to ascertain whether similarities exist.

Experimental groups #2-#5 were combined in order to observe more general trends in the data. Experimental group #1 was not used, since it contained no error cluster; and experimental group #6 was not considered because it involved repetition of a whole word rather than of an initial cluster or syllable. Groups #1 and #6 also contained small numbers of observations and were thus less likely to affect group trends.

As mentioned previously, subjects were told to produce the experimental utterances at their fastest speaking rate. Assuming that experimental utterances were indeed produced at a "fastest" rate of speech, one might further assume that segmental and cluster length would decrease; however, this is not the case, as can be seen in Tables 3 and 8. One explanation for this could be that only vowels, resonants and perhaps pauses make a difference to rate, while stops and fricatives are only slightly affected, if at all. In addition, a possible explanation for this subjectively faster rate of speech could be the change in duration of the hesitation pause and the relation of semantic content to pausal time.

As Goldman-Eisler (1968) reports, variations in the overall rate of speech or an increase in rate were found to be variations in the amount of pausing. She concludes that rate of speech based solely on articulatory activity remained relatively invariant.

Goldman-Eisler (1968) also observes trends in pausal time in relation to semantic content. When subjects interpreted meaning, pausal time was twice as great as when they described content. She examined this phenomenon with respect to degrees of spontaneity in speech. Where semantic content becomes less and less a factor in speech, as in repetition of the same utterance, she found that there was a decline in pausal time after the first repetition and a further decrease in subsequent repetitions.

Considering these interpretations, one would not assume that cluster length would decrease with an increase in rate but would remain approximately the same, attaining its shortest duration in the last (i.e., fifteenth) repetition where semantic content is most familiar.

General results can be grouped for discussion purposes: The longest sibilants are those before the velar /k/, while the longest stop consonant is the bilabial /p/ and the shortest the alveolar /t/. In the control group data, the stop consonant was considered to determine the length of the cluster. Hypothesizing that such a constraint holds for the experimental group data, one might expect the corrected /sp/-clusters to be the longest and the corrected /st/-clusters to be the shortest. The results presented in Table 9 support such an hypothesis.

A possible physiological explanation for the finding that /sp/ and /sk/ are longer in duration than /st/ would be that the former involve slower moving articulatory musculature (e.g., lips and body of the tongue) for the stop production, while both /s/ and /t/ involve faster moving, more highly innervated tongue tip musculature. Furthermore, /t/ is homorganic with /s/ (i.e., place of articulation is the same, only the manner differs), whereas the heterorganic clusters /sp/ and /sk/ require several different muscles in order to complete the articulatory gestures. These different movements should tend (logically) to make production slower for /sp/ and /sk/ than for /st/. Haggard (1973) also discusses abbreviation in certain homorganic clusters: He reports varied individual differences and supposes that durations can be controlled by oral pressure feedback.

Perhaps the most intriguing finding of the present investigation is that the elements of an error cluster are always longer than those of the repetition, or correction production. It is of interest to note the possible effects of stress and following vowel environment on cluster duration.

Of the thirty-six test words embedded in the three "tongue-twisters", all but two had primary stress on the first syllable containing the cluster. Word stress was therefore not considered to be a major factor in determining cluster length. Similarly, the following vowel environment was analyzed to determine whether repetition errors occurred more frequently before some vowels and not others. Of 118 repetition errors produced by the six subjects, 39 errors occurred before the vowel /æ/, 29 before /ɩ/, 24 before /a/, 9 before /ɛ/, 5 before /oɷ/, 5 before /i/, 4 before /ɔ/, and 3 before /eɩ/. More repetition errors occurred before the vowels /æ/, /ɩ/ and /a/ than before the others. Since data were not analyzed to take following vowels into account, no explanation as to their significance can be offered at this time. The import of these results are the topic of the next section.

5.3 Theoretical Considerations

Results from the experimental group data strongly suggest that duration of a segment or cluster already produced can affect subsequent articulations. It seems that excessive duration of the cluster as a whole or of either of its component parts may violate some sort of timing constraint on the system in the production of a given utterance; this violation causes the production to be halted in mid-word, a recalibration to be effected, and a correction to be produced.

If, however, an error is made with respect to the phonetic form of the utterance (such as a substitution error), the timing constraint hypothesized above may not be violated, and the utterance would not have to be repeated. At present it is not known whether addition or omission errors violate such a timing constraint in words, or if other segments within words lengthen or shorten to make room for an extra element or to fill up an empty space. If such a view is tenable, the hesitation errors (where an intra-word pause or prolongation of a segment occurs) can be considered to be a step earlier than a repetition error; i.e., within certain limits, segmental prolongation or pause insertion will be enough for the system to recalibrate, but if such recalibration does not take place quickly enough, then production is completely halted, with the utterance being reproduced, yielding what has been termed here a repetition error.

Consideration solely of the repetition errors does not allow for determination of whether (a) a delay in production of the next phoneme caused the repetition, (b) the repetition is a correction of an excessively long segment, or (c) the repetition is a correction of an excessively long cluster just uttered. A combination of (a), (b) and (c) is a likely solution; i.e., excessively long segments in a cluster are produced which causes a delay in production of the next phoneme, and the second production in a repetition error is a correction of the timing violation. In support of this solution, let us again consider hesitation errors: As mentioned previously, no repetition occurs in such errors (perhaps) because (a) the duration of the hesitation is not long enough to violate a timing constraint, and thus (b) the phoneme transitions have not been lost so that production of the word can be continued. This suggests that when a repetition takes place, it is due to excessive duration which causes a delay violating a timing constraint and results in loss of phoneme transitions. Consequently, the utterance cannot be continued, and a correction of the excessive duration of the first production is ordered. If such a notion is adopted, then how might the system determine that a timing constraint has been violated and a correction called for?

In order to accomplish correction of an element previously uttered, the system must first know that the element was incorrect and what the target should have been; i.e., feedback must be present in order for the system to recognize the duration error, and comparison of output with the originally planned target element must occur in order for a correction to be executed.

The "analysis-by-synthesis" model of speech perception and production (cf. Bell et al., 1961) incorporates a feedback method of speech spectrum analysis by which correction of an element previously uttered could be accomplished. In terms of this model, the "spectrum generator" produces output comparable to stored speech data (i.e., the first production of the utterance). The "comparator" then computes the difference between the input speech spectra it has received via a feedback loop and the original target utterance just generated. Trial spectra are synthesized by the "strategy component" until minimum error is obtained in matching and in this case a correction of a previous error utterance is generated.

On the physiological level, feedback can occur via the acoustic and/or proprioceptive channels (as discussed in Sections 1.31 and 1.32), i.e., via bone and air conduction and/or the gamma motor system. In the present study there is no way to determine if both channels are in use at all times, or if there is intermittent monitoring by one or both channels during speech (as suggested earlier). One argument for intermittent feedback during speech is that omission and addition errors are usually not corrected. One may speculate that such errors do not violate a timing constraint placed on the word, but it may be more likely that such errors have not been "caught" by the system due to intermittent monitoring. In any case, the speaker/hearer receives feedback, and some comparison with the original target must take place -- facts which models of speech production (and perception) should account for.

The notions discussed thus far may be viewed in light of Fairbanks's (1954) interpretation of the speech production system as a servosystem. He suggested that there is continuous monitoring via the acoustic mode, by which we compare output to input and thereby manipulate production. Monitoring solely by means of the auditory channel is most unlikely, since an adventitiously deafened person does not lose his speech immediately after an injury (as noted in Section 1.62); perhaps such an individual can rely on his proprioceptive feedback from the articulators to supplement barely discernible auditory signals.

For the normal hearing person, Abbs (1973) proposed a "variable" servosystem (cf. Section 1.63), a model which closely acquaints proprioceptive feedback and input/output comparisons. This speech production system is efficient, since it employs feedback depending on its requirements, via the gamma or spindle motor system. In order to consider which system might be in operation, let us speculate how repetition errors might violate timing constraints.

The violation of a timing constraint could be due to a "lapse" in continuous auditory feedback because of a faster rate of speech; i.e., the auditory monitoring system (and "comparator") lag behind production to the extent that a lengthening of segments occurs which, in the case of repetition errors, violates a timing constraint. Perhaps for this reason these errors closely resemble the repetition errors produced under the influence of delayed auditory feedback, where a delay results in such a "lapse" and a repetition is produced (cf. Fairbanks & Guttman, 1958; Lee, 1951). If such a lapse does occur, then repetition errors are not just a production problem, but also a perception problem. In relating this notion of auditory feedback's lagging speech production to the "analysis-by-synthesis" model, we might speculate that it is the "comparator", in computing the difference between input and output, which has caused the delay, and a recalibration is necessary for it to catch up to production. The delay thus violates a timing constraint here too, and a corrected utterance must be produced.

To account for the fact that not all speech errors are corrected, one may speculate that auditory feedback monitors only general sound patterns and intonation, rhythm and stress patterns, while proprioceptive feedback monitors intermittently for phonetic deviancies. In accordance with Abbs (1973) this intermittent proprioceptive feedback concerns the gamma or spindle motor system, which (1) maintains length or rate of change of length of a muscle, (2) initiates contraction, (3) provides damping movements to prevent overshoot, and (4) supplements auditory feedback when necessary.

There is a reasonable basis for supposing that the gamma motor system is operational with respect to phonetically deviant repetition errors (such as those observed in this study). If we consider the shortest pause between repetitions observed, here approximately 20 ms, and the "turn around time" or delay for operation of the spindle system, ca. 20-80 ms, we find that the two figures overlap, and the interpretation is not contradicted.

All told, a model of speech production must account for normal utterances, as well as for speech errors. Moreover, it must allow for appropriate types of feedback and be an efficient system for relating speech production to perception. The variable servomonitor system outlined above incorporates both continuous auditory feedback and intermittent proprioceptive feedback, which are used in perceiving input and using it to manipulate output. This system also provides a plausible account of the speech errors and of their production as described in this study. Theoretically, the system advocated here provides an efficient means for producing, monitoring and correcting speech production.

5.4 Limitations of the Present Study

One of the purposes of this investigation was to examine speech errors under conditions of rapid repetition of three "tongue-twisters". It was hypothesized that errors obtained using a faster than normal speaking rate might be due to at least the following: (1) The words were similar in phonological form and phonetic content (i.e., twelve occurrences each of /sp/, /st/ and /sk/ in word-initial position), and (2) the faster rate of speech which, when combined with (1), results in error productions. These involve variables such as psychological and physical stress and fatigue which were not controlled for -- if indeed they can be controlled for -- in this study.

The following limitations apply to any interpretation of the data:

(1) Generation of speech errors may not produce the same type of errors as those produced in spontaneous speech. They may be due to memory limitations, which, when combined with speaking rate, put undue stress on the speaker, who may then produce "unnatural" errors. As such, the errors described may be artifacts of the experimental method employed.

(2) Words with initial fricative plus stop consonant clusters were not controlled for number of syllables, word class, stress placement or place in sentence.

The above were considered to be the major limitations of this study, and due consideration for the control of such variables should be given to future research in this area.

5.5 Summary and Conclusions

The present study first examined the means by which the generation and the classification of speech errors could be accomplished. It was found that a "tongue-twister", which contained many occurrences of words with similar phonological form (i.e., with word-initial /sp/, /st/ and /sk/ clusters), produced at a subject's fastest speaking rate and repeated many times would generate the most speech errors. These errors were then classified according to categories combined from the literature on speech errors and on DAF research. Fifty percent of all errors produced by subjects in this investigation were of the repetition type. Consideration of the possible causes of the errors encountered led to the detailed examination of the word-initial cluster and component segment durations.

The experimental investigation yielded the following general results:

(1) The stop consonant in a given cluster seems to determine the overall cluster duration, since the duration of /s/ irrespective of context remains fairly constant.

(2) The clusters /sp/ and /sk/ are longer in duration than /st/, which may be attributable to the slower moving articulatory musculature associated with /p/ and /k/ production compared with the faster moving, more highly innervated tongue tip musculature involved in the production of /s/ and /t/.

(3) The cluster segments in the error productions were consistently longer in duration than in the second, or corrected, production (which approximated more closely the values obtained for the control group data than might have been otherwise expected due to methodological differences).

In light of the above results, it was speculated that the excessive duration of the cluster (or of its component parts) violated a timing constraint on the production of an utterance, whereupon phoneme transitions are lost to the system, following which a recalibration must take place and a correction produced. From such considerations it was inferred that feedback must be present in order for the system to recognize the duration error, to compare it with planned output, and finally to execute a correction.

On the physiological level, feedback was considered to be both continuous (auditory channel) and intermittent (proprioceptive channel, involving the gamma motor system); the latter may supplement auditory feedback and scan for deviant phonetic elements, while the former monitors general sound patterns, particularly suprasegmental patterns.

As a result of these considerations, it was hypothesized that a timing constraint is imposed by the system. When this constraint is violated, perhaps due to a delay in auditory feedback processing occasioned by the faster than normal speaking rate, a speech error occurs. Because of this, repetition errors were regarded as much a perception problem as a production problem.

REFERENCES

Abbs, J.H. 1973. The influence of the gamma motor system on jaw movements during speech: A theoretical framework and some preliminary observations. Journal of Speech and Hearing Research 16. 175-200.

Alston, W., R.W. Angel, F.S. Fink, and W.W. Hofmann. 1967. Motor activity following the silent period in human muscle. Journal of Physiology 190. 189-202. [Cited in Abbs, 1973.]

Bell, C.G., H. Fujisaki, J.M. Heinz, K.N. Stevens, and A.S. House. 1961. Reduction of speech spectra by analysis-by-synthesis techniques. Journal of the Acoustical Society of America 33. 1725-1736.

Bernstein, N.A. 1967. The Coordination and Regulation of Movements. Oxford: Pergamon Press. [Cited in Ohala, 1973.]

Berry, M., and J. Eisenson. 1956. Speech Disorders: Principles and Practices of Therapy. New York: Appleton-Century-Crofts.

Black, J.W. 1951. The effect of delayed side-tone upon vocal rate and intensity. Journal of Speech and Hearing Disorders 16. 56-60.

Boomer, D.S., and J.D.M. Laver. 1968. Slips of the tongue. British Journal of Disorders of Communication 3. 1.12.

Borden, G.J., and T. Gay. 1975. Durations of articulator movements for /s/-stop clusters. Haskins Laboratories Status Report on Speech Research SR-44. 147-160.

Bowman, J.P. 1971. The Muscle Spindle and Neural Control of the Tongue. Springfield, Ill.: Charles C Thomas.

Campbell, E.J.M. 1968. The respiratory muscles. Sound Production in Man, ed. by A. Bouhuys, 135-140. New York: New York Academy of Sciences. [Cited in Abbs, 1973.]

Chase, R.A. 1958. Effect of delayed auditory feedback on the repetition of speech sounds. Journal of Speech and Hearing Disorders 23. 583-590.

Euler, C. von. 1966. Proprioceptive control in respiration. Muscular Afferents and Motor Control, ed. by R. Granit, 197-208. New York: Wiley. [Cited in Abbs, 1973.]

Fairbanks, G. 1954. Systematic research in experimental phonetics: 1. A theory of the speech mechanism as a servosystem. Journal of Speech and Hearing Disorders 19. 133-139.

Fairbanks, G. 1960. Voice and Articulation Drillbook, 2nd edition. New York: Harper & Row.

Fairbanks, G., and N. Guttman. 1958. Effects of delayed auditory feedback upon articulation. Journal of Speech and Hearing Research 1. 12-22.

Freud, S. 1924. Zur Psychopathologie des Alltaglebens, 10th edition. Leipzig: Internationaler Psychoanalytischer Verlag. [Cited in Fromkin, 1971.]

Fromkin, V.A. 1968. Speculations on performance models. Journal of LInguistics 4. 47-68.

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

Fromkin, V.A. 1973. Slips of the tongue. Scientific American 229 (6). 110-117.

Goldman-Eisler, F. 1968. Psycholinguistics: Experiments in Spontaneous Speech. London: Academic Press.

Grinker, R.R., and P.C. Bucy. 1949. Neurology. Springfield, Ill.: Charles C Thomas.

Haggard, M. 1973. Abbreviation of consonants in English pre- and post-vocalic clusters. Journal of Phonetics 1. 9-24.

Hardy, J.C. 1970. Development of neuromuscular systems underlying speech production. Speech and the Dentofacial Complex: The State of the Art (ASHA Reports No. 5), 49-68. Washington, D.C.: American Speech and Hearing Association.

Johnston, T.B., and J. Whillis (Eds.). 1954. Gray's Anatomy -- Descriptive and Applied, 31st edition. London: Longmans, Green and Co.

Kent, R.D., and K.L. Moll. 1975. Articulatory timing in selected consonant sequences. Brain and Language 2. 304-323.

Konigsmark, B.W. 1970. Neuroanatomy of speech. Speech and the Dentofacial Complex: The State of the Art (ASHA Reports No. 5), 3-19. Washington, D.C.: American Speech and Hearing Association.

Kozhevnikov, V.A., and L.A. Chistovich. 1965. Speech: Articulation and Perception. [Joint Publications Research Service, No. 30,543.] Washington, D.C.: U.S. Department of Commerce.

Lee, B.S. 1950. Effects of delayed speech feedback. Journal of the Acoustical Society of America 22. 824-826.

Lee, B.S. 1951. Artificial stutter. Journal of Speech and Hearing Disorders 16. 53-55.

Lehiste, I. 1971. Temporal organization of spoken language. Form & Substance, ed. by L.L. Hammerich, R. Jakobson, and E. Zwirner, 159-169. Copenhagen: Akademisk Forlag.

Lehiste, I., and G.E. Peterson. 1959. Vowel amplitude and phonemic stress in American English. Journal of the Acoustical Society of America 31. 428-435.

Lenneberg, E.H. 1967. Biological Foundations of Language. New York: Wiley.

MacKay, D.G. 1970. Spoonerisms: The structure of error in the serial ordering of speech. Neuropsychologia 8. 323-350.

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

MacNeilage, P.F. 1972. Speech physiology. Speech and Cortical Functioning, ed. by J. Gilbert, 1-72. New York: Academic Press.

Meringer, K. 1908. Aus dem Leben der Sprache. Berlin. [Cited in Fromkin, 1971.]

Meringer, K., and K. Mayer. 1895. Versprechen und Verlesen, Eine psychologisch linguistische Studie. Vienna. [Cited in Fromkin, 1971.]

Merton, P.A. 1953. Speculations on the servo-control of movement. The Spinal Cord, ed. by G.E.W. Wolstenholme, 247-255. London: Churchill. [Cited in Abbs, 1973.]

Mortimer, E.M., and K. Akert. 1961. Cortical control and representation of fusiform neurons. American Journal of Physical Medicine 48. 228-248. [Cited in Abbs, 1973.]

Netter, F.H. 1972. The CIBA Collection of Medical Illustrations, Volume 1: Nervous System. Summit, N.J.: CIBA Pharmaceutical Co.

Nooteboom, S.G. 1969. The tongue slips into patterns. Nomen: Leyden Studies in Linguistics and Phonetics, ed. by A.G. Sciarone et al., 114-132. The Hague: Mouton.

Ohala, J.J. 1970. Aspects of the control and production of speech. Working Papers in Phonetics, No. 15. Los Angeles: University of California.

Ohala, J.J. 1973. The temporal regulation of speech. Collection of Papers Presented at Symposium on Auditory Analysis and Perception of Speech (Leningrad, August 21-24), ed. by H. Fujisaki, 516-538. Tokyo: University of Tokyo.

Penfield, W., and T. Rasmussen. 1950. The Cerebral Cortex of Man. New York: Macmillan.

Penfield, W., and L. Roberts. 1959. Speech and Brain-Mechanisms. Princeton, N.J.: Princeton University Press.

Ruch, T.C., H.D. Patton, J.W. Woodbury, and A.L. Towe. 1967. Neurophysiology. Philadelphia: W.B. Saunders. [Cited in Abbs, 1973.]

Sataloff, J. 1966. Hearing Loss. Philadelphia: Lippincott.

Schwartz, M.F. 1970. Duration of /s/ in /s/-plosive blends. Journal of the Acoustical Society of America 47. 1143-1144.

Sears, T.A., and J. Newsome Davis. 1968. The control of respiratory muscles during voluntary breathing. Sound Production in Man, ed. by A. Bouhuys, 183-190. New York: New York Academy of Sciences. [Cited in Abbs, 1973.]

Smith, J.L. 1969. Fusiform Neuron Block and Voluntary Arm Movement in Man. Unpublished doctoral dissertation, University of Wisconsin. [Cited in Abbs, 1973.]

Smith, T.S. 1973. Review of Bowman (1971). Journal of Phonetics 1. 171-179.

Stark, L. 1968. Neurological control systems: Studies in bioengineering. The Hand, 301-410. New York: Plenum. [Cited in Abbs, 1973.]

Stromsta, C. 1962. Delays associated with certain sidetone pathways. Journal of the Acoustical Society of America 34. 392-396. [Cited in Van Riper, 1971.]

Sturtevant, E.H. 1917. Linguistic Change. Chicago: University of Chicago Press.

Sturtevant, E.H. 1947. An Introduction to Linguistic Science. New Haven: Yale University Press.

Van Riper, C. 1971. The Nature of Stuttering. Englewood Cliffs, N.J.: Prentice-Hall.

APPENDIX A

The first paragraph of "The Rainbow Passage" (Fairbanks, 1960, p. 127):

"When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow."

APPENDIX B

Paragraphs with embedded word-initial /sC/-clusters:

/sp-/: "In school I had a friend nicknamed 'Spud', who was terrible at most everything. One afternoon in the school spelling bee, Spud spelt 'spaghetti' speedily and won the school spelling prize."

/st-/: "In school I also had two friends, Stan and Stewart, who were always fighting with one another. One day while in a fight, Stewart stepped on Stan's stomach, and he died three days later."

/sk-/: "Another friend from my school days was Skana. She liked doing anything better than going to school. In fact, Skana skipped school so skillfully that no one knew what happened to her."

APPENDIX C

"Tongue-twisters" with embedded word-initial /sC/-clusters:

/sp-/: "A spectre of a spirited spectacled Spanish 'Spartan' called Spinoza ate the spotted spiced spinach sporadically with a spoon before it spoiled."

/st-/: "The stalwart stallion and the statuesque steer were startled by the stoical stable staff stealing stagnant stew off the stove."

/sk-/: "Scandinavians skillfully skin and scald scallions, scallops, scampies and skimpy scorpions until they are scarcely scarlet, then scarf them down."

APPENDIX D

PHONETIC TRANSCRIPTIONS AND CODING

The repetition errors for each subject in this study are listed according to type (or experimental group), phonetic transcription (using a modified version of the International Phonetic Alphabet), and standard orthographic form of the target utterance. All errors involve the word-initial clusters: /sp-/, /st-/, /sk-/.

SUBJECT 1. REPETITION ERRORS. (Total: 18)

Experimental Group #1: (Phoneme Repetition)	/sp-/	[s?spɪnoozə]	'Spinoza'
	/st-/	[s?stooɪkl̩]	'stoical'
Experimental Group #2: (Cluster Repetition, with a Pause)	/sk-/	[sk?skæmpiz]	'scampies'
		[skʰ?skɪlfoli]	'skillfully'
		[skʰ?skæmpiz]	'scampies'
		[sk?skɪlfoli]	'skillfully'
Experimental Group #3: (Cluster Repetition, without a Pause)	/sp-/	[spspɔɹædɪkli]	'sporadically'
	/st-/	[stʰstætʃjuɛsk]	'statuesque'
	/sk-/	[skskɪlfoli]	'skillfully'
Experimental Group #4: (/sCV-/ Syllable Repetition)	/sp-/	[spəˆ?spadɪd]	'spotted'
		[spɛ?spɛktʰəʴ]	'spectre'
		[spʰəɪ?spadɪd]	'spotted'
	/st-/	[stʰə?staɹtl̩d]	'startled'
	/sk-/	[skʰa?skaɹlɛt]	'scarlet'
Experimental Group #5: (/sCVC-/ Syllable Repetition)	/st-/	[stʰæks?stætʃjuɛsk]	'statuesque'
		[stæt?stætʃjuɛsk]	'statuesque'
	/sk-/	[skaɹ?skaɹlɛt]	'scarlet'
		[skæm?skaɹf]	'scarf'
Experimental Group #6: (Word Repetition)	No observations.		

SUBJECT 2. REPETITION ERRORS. (Total: 28)

Experimental Group #1: (Phoneme Repetition)	/sp-/	[s?spɛktəkəld]	'spectacled'
	/st-/	[s?stætʃuɛsk]	'statuesque'
	/sk-/	[s?skæljənz]	'scallions'
		[s?skɩmpi]	'skimpy'
		[s?skʰaɹlɛt]	'scarlet'
		[s?skæləps]	'scallops'
		[s?skɩmpi]	'skimpy'
		[s?skæljənz]	'scallions'
		[s?skæmpiz]	'scampies'
Experimental Group #2: (Cluster Repetition, with a Pause)	/sp-/	[s˺pʰ?spɔɹædɩkli]	'sporadically'
	/sk-/	[skʰ?skɩlfoli]	'skillfully'
		[sk?skæljənz]	'scallions'
Experimental Group #3: (Cluster Repetition, without a Pause)	/st-/	[ststʰilɩŋ]	'stealing'
	/sk-/	[sk:skɩn]	'skin'
Experimental Group #4: (/sCV-/ Syllable Repetition)	/sp-/	[spʰəɩ?spadɩd]	'spotted'
	/st-/	[stʰæ?stætʃuɛsk]	'statuesque'
		[stɛstæljən]	'stallion'
		[stʰo?stoɷɩkl̩]	'stoical'
	/sk-/	[stɩ?skɩlfoli]	'skillfully'
Experimental Group #5: (/sCVC-/ Syllable Repetition)	/sp-/	[spʰiɹ?spʰiɹɩdɩd]	'spirited'
		[spɛkʰ?spɛktəʴ]	'spectre'
	/st-/	[stʰoɷɩkstʰoɷɩkʰl̩]	'stoical'
	/sk-/	[stɩlf?skɩlfoli]	'skillfully'
		[skɩmp?skɩmpi]	'skimpy'
		[skæmpʰ?skæmpiz]	'scampies'
		[skæntʰ?skæmpiz]	'scampies'
		[skaɹ?skɛɩɹsli]	'scarcely'
Experimental Group #6: (Word Repetition)	/sp-/	[spɛkl̩d?spɛktʰækl̩d]	'spectacled'

SUBJECT 3. REPETITION ERRORS. (Total: 15)

Experimental Group #1: (Phoneme Repetition)	/sk-/	[sʔskɪmpʰi]	'skimpy'
		[sʔskɪmpi]	'skimpy'
		[sʔskæləps]	'scallops'
		[sʔskʰɪmpi]	'skimpy'
		[sʔskaɹf]	'scarf'
Experimental Group #2: (Cluster Repetition, with a Pause)	/sk-/	[skʰʔskændəneɪvijənz]	'Scandinavians'
		[skʰʔskʰæmpiz]	'scampies'
Experimental Group #3: (Cluster Repetition, without a Pause)	/sk-/	[sk:ʰskaɹf]	'scarf'
		[sk:skɪlfʊli]	'skillfully'
Experimental Group #4: (/sCV-/ Syllable Repetition)	/sk-/	[skʰɪʔskʰɪmpi]	'skimpy'
		[skʰaʔskʰaɹlɛt]	'scarlet'
Experimental Group #5: (/sCVC-/ Syllable Repetition)	/st-/	[stægʔstætʃuɛsk]	'statuesque'
Experimental Group #6: (Word Repetition)	/sk-/	[skɔɹpizʔskæləps]	'scallops'
		[skʰɪnʔskaɹf]	'scarf'
		[skʰɪnʔskʰɪn]	'skin'

SUBJECT 4. REPETITION ERRORS. (Total: 33)

Experimental Group #1:	/sp-/	[s?spɛktəkʰl̩d]	'spectacled'
(Phoneme Repetition)	/st-/	[s?stebl̩]	'stable'
		[s?stæljən]	'stallion'
Experimental Group #2:	/st-/	[stʰ?steɪbl̩]	'stable'
(Cluster Repetition, with a Pause)		[stʰ?stalwəʴt]	'stalwart'
		[stʰ?stoʊɪkʰəl]	'stoical'
		[stʰ?stægn̩t]	'stagnant'
		[stʰ?stʰilɪŋ]	'stealing'
		[skʰ?skɪn]	'skin'
		[sk:ʰskɪmpi]	'skimpy'
		[sk?skɔɹpijənz]	'scorpions'
		[sk:?skɪmpi]	'skimpy'
		[skʰ?skæljənz]	'scallions'
Experimental Group #3:	/sp-/	[sp:spʰɔɹædɪkli]	'sporadically'
(Cluster Repetition, without a Pause)		[sp:ʰspadɪd]	'spotted'
		[sp:spaɹ?n̩]	'Spartan'
	/sk-/	[sk:ʰskɪlfoli]	'skillfully'
Experimental Group #4:	/sp-/	[spa?spænɪʃ]	'Spanish'
(/sCV-/ Syllable Repetition)	/st-/	[stæ?stʰæljən]	'stallion'
		[ste?steɪbl̩]	'stable'
		[stæ?stætʃuɛsk]	'statuesque'
		[stɪ?stoʊv]	'stove'
		[stæ?stægn̩t]	'stagnant'
		[sta?stalwəʴt]	'stalwart'
	/sk-/	[skɛ^?skæljənz]	'scallions'
Experimental Group #5:	/sp-/	[spad?spiɹɪdɪd]	'spirited'
(/sCVC-/ Syllable Repetition)	/st-/	[stægn̩?stætʃuɛsk]	'statuesque'
		[staljəʴ?stalwəʴt]	'stalwart'
	/sk-/	[skɪl?skɪmpi]	'skimpy'
		[skɪl?skɪn]	'skin'
Experimental Group #6:	/sk-/	[skʰaɹfɛt?skʰaɹlɛt]	'scarlet'
(Word Repetition)		[skʰɪmpʰə?skɪmpi]	'skimpy'
		[skɪlf?skɪmpi]	'skimpy'

SUBJECT 5. REPETITION ERRORS. (Total: 11)

Experimental Group #1:	No observations.		
Experimental Group #2:	No observations.		
Experimental Group #3: (Cluster Repetition, without a Pause)	/sk-/	[sk:skæləps]	'scallops'
Experimental Group #4: (/sCV-/ Syllable Repetition)	/sp-/	[spʌ?spadɪd]	'spotted'
	/st-/	[stʰoo?stʰætʃjuɛsk]	'statuesque'
	/sk-/	[skʰæ?skæləps]	'scallops'
		[skæ?skæləps]	'scallops'
		[skʰɪ?skʰɪn]	'skin'
Experimental Group #5: (/sCVC-/ Syllable Repetition)	/sk-/	[skʰæm?skæləps]	'scallops'
		[skæl?skændɪnevijənz]	'Scandinavians'
		[staɹ?skʰaɹlɛt]	'scarlet'
Experimental Group #6: (Word Repetition)	/sp-/	[spɛktʰə?spɛktʰəɹ]	'spectre'
	/sk-/	[stʰɛɹslіskʰɛɪɹsli]	'scarcely'

SUBJECT 6. REPETITION ERRORS. (Total: 13)

Experimental Group #1:	No observations.		
Experimental Group #2: (Cluster Repetition, with a Pause)	/sp-/	[spʔspiɹɪdɪd]	'spirited'
		[spʰʔspɛktʰəkəld]	'spectacled'
		[spʰʔspænɪʃ]	'Spanish'
		[spʰʔspaɹtn̩]	'Spartan'
	/sk-/	[skʰʔskaləps]	'scallops'
		[skʰʔskɪn]	'skin'
		[skʰʔskɪlfoli]	'skillfully'
Experimental Group #3: (Cluster Repetition, without a Pause)	/st-/	[st:stætʃjuɛsk]	'statuesque'
Experimental Group #4: (/sCV-/ Syllable Repetition)	/st-/	[stəʔstalwəʴt]	'stalwart'
	/sk-/	[skʰɪʔskɪmpi]	'skimpy'
		[skaʔskaləps]	'scallops'
		[skæʔskæljənz]	'scallions'
Experimental Group #5: (/sCVC-/ Syllable Repetition)	/st-/	[stægʔstægnənt]	'stagnant'
Experimental Group #6:	No observations.		