DURATIONS OF ACOUSTIC SEGMENTS UNDER SYNCHRONOUS AND DELAYED FEEDBACK CONDITIONS

by

Ingrid Jeffrey

B.A., Simon Fraser University, 1972 .

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Department

of

Paediatrics

Division of Audiology and Speech Sciences

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July,1974

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.

Department of Paediatrics

The University of British Columbia
Vancouver 8, Canada

Date October 9th, 1974

ABSTRACT

The present study investigates durations of acoustic segments under synchronous and delayed feedback conditions.

Three subjects read a passage and four sentences at their normal rate (NORMAL), at a slow rate (SLOW SAF) under synchronous feedback, and then at a slow rate (SLOW DAF) and at a maximally fast rate (FAST DAF) under delayed auditory feedback. The delay used was determined for each subject so it would produce maximum speech disturbance.

The hypotheses under test, on the basis of a pilot study were: (a) that under SLOW SAF and under delayed feedback, vowels would be prolonged proportionately more than consonants, and that continuants would be prolonged proportionately more than obstruents; and (b) that close vowels would be prolonged more under SLOW SAF and open vowels more under DAF conditions. Position in the syllable was expected to affect selectively increases in duration under DAF.

The results, based on normalized data, confirmed that vowels were proportionately more prolonged than consonants under SLOW SAF and DAF conditions. Continuant consonants were proportionately more prolonged than obstruents under DAF conditions but not consistently under SLOW SAF. For the obstruents prevocalic closure was consistently more prolonged under DAF than under SLOW SAF.

The hypothesis that close vowels would be prolonged

more under SLOW SAF and open vowels more under DAF was not in general confirmed. The study further indicated that position in the syllable not only affected durations of segments under DAF but also under SLOW SAF: Under SLOW SAF consonants in postvocalic position were proportionately more prolonged and under DAF consonants in prevocalic position were proportionately more prolonged.

It was also found that the durations of vowels in words such as function words increased proportionately more than the duration of vowels in other words and that the duration of the vowel /u/ increased proportionately more than that of open vowels. Moreover, /u/ occurred more frequently than other vowels in phonetic environments in which vowels are normally of longer duration.

Finally, it is hypothesized that the syllable is a monitored unit under DAF and that the reflex level is involved in the DAF-induced speech disturbances.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I would like to express my sincere appreciation for the help I have received in completing this thesis. My thanks go especially

- To Dr. A-P. Benguerel without whose time and advice this thesis would not have been possible;
- To Dr. J.B. Delack, not only for being on my thesis committee, but particularly for his concern and interest always;
- To Dr. R. Gregg for his time spent serving on my committee;
- To Dr. J.H. Gilbert and Carolyn for their kind encouragement;
- To Dr. R.P. Gannon for permission to use the facilities of the Audio-Vestibular Unit at Vancouver General Hospital;
- To The Psychology Department at Simon Fraser University for the loan of the Körting tape recorder;
- To Jim and Kathy for their love and understanding;
- To Sharon, Pat, Lynne, Betty and Meralin for the closeness we shared;
- To Robin for her work on the drawings;
- To Rose and Margaret for typing;
- And to my subjects for their patience and cooperation.

Chapter 1

INTRODUCTION

The importance of sensory feedback for the smooth execution of voluntary movements has long been recognized. In the production of speech a number of voluntary movements involving the various articulators, the respiratory mechanism, and the vocal folds have to be precisely coordinated in time. To add to the complexity of such a set of coordinated movements, the neural commands for the various movements required for the production of a sound cannot always be initiated at the same point in time. The sequence at which impulses will be fired is determined by a number of factors such as the travelling time required by the impulse to reach its destination, "mass of tissue that has to be moved and the amount of work that must be performed" (Lenneberg, 1967, p. 100). Furthermore, during the production of a particular sound, features of later sounds may already be present. Speech, in other words, does not consist of one discrete set of coordinated movements resulting in the production of a sound the length of a phoneme followed by another, but consists of a set of discrete neural commands with associated movements coordinated in time. The length of the production unit or units in which movements are integrated has been and continues to be a subject of investigation.

It has been suggested (Smith, 1962) that the

motion patterns of speech are controlled and monitored by a system of feedback mechanisms. With tactile, kinesthetic, bone- and air-conduction feedback channels available to such a system, what is the precise nature of these feedback channels, their utilization in the process of monitoring and controlling, and their relative importance?

By means of interference with the feedback channels, investigators have sought answers to these questions. A considerable amount of data has been collected about the auditory feedback mechanism. For example, it has been found that when both bone- and air-conducted auditory feedback are masked by high intensity noise, vowels are considerably affected both in length and quality, and control of the soft palate is lost to some extent. It has also been observed that pitch increases and that changes of intonation patterns and voice occur under this condition (Ladefoged, 1967,pp. 163-164).

Introducing a delay between the speaker's vocal output and his air-conducted auditory input, a method known as delayed auditory feedback (DAF), also referred to as delayed side-tone, is a frequently employed method of interference, delaying the usual time of arrival of the auditory feedback during speech production. Depending on the occluding properties of the headset earmuffs, the sound absorbing properties of the environment, the length of the delay, and the amplification of the delayed signal, fragments of undelayed auditory feedback will be available to the speaker.

Early investigators (Lee, 1950a, 1950b, 1951; Black, 1951; Fairbanks, 1955) found that DAF slows the rate of speech production, increases vocal intensity and fundamental frequency, and results in speech errors such as substitutions, omissions and additions; the latter consist primarily of repetitions (Fairbanks and Guttman, 1958). Such DAF-induced repetitions have also been termed "artificial stutter" (Lee, 1951).

The delay has been varied experimentally, and maximal speech disturbance has been reported at 180 msec (Black, 1951), and at 200 msec (Fairbanks, 1955). Still another study reports 180 msec to be the maximally disturbing delay for males and 270 msec for females (Mahaffey and Stromsta, 1965). It might be mentioned that considerable individual differences exist in the ability to cope with delayed feedback as well as in the strategies employed under such conditions. Considerable variation from subject to subject in the resulting behaviour can be expected. Between-subject comparisons must therefore be viewed with caution, particularly since the measures employed in most of the past DAF studies are of a very global nature. Measures of total reading time in relation to speech units such as phonemes, syllables, or words contained in the read passage implicitly either assume a uniform slowing of all elements in the test passage or a slowing mainly due to increase of articulatory errors. Only what Allen (1973) has termed global rate, that is, the average rate at which

phonemes, syllables, and words are produced, can be accounted for by such measures. A more detailed analysis is needed to learn anything about what Allen referred to as local rate, that is, "the specification of segment durations within syllables" (ibid., p. 222).

As Allen has pointed out, local rate is partly a function of global rate. How then does local rate change as a function of global rate? Linear relationships can be expected only for a narrow range of tempos (ibid., p. 222).

This study proposes to investigate changes in local rate as a function of global rate, under synchronous feedback (SAF) as well as under DAF. Only if global rate under SAF is altered and the effects on local rate are determined, can the effects of DAF-altered global rate on local rate be assessed.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

In this section the pertinent DAF experiments will be briefly described in chronological order. The results and underlying assumptions of the experimenters will be discussed in an attempt at a synthesis of available information as it relates to the DAF mechanism and possible units of speech production.

2.2 Experiments

A summary in table form of the experiments to be described in the following pages can be found in the appendix (Appendix A).

Lee (1950a), one of the earliest DAF experimenters, first called the method of DAF to the attention of investigators as a means of studying feedback processes. In his own experiments he found that some subjects were able to avoid to some extent the effects of DAF, presumably as a result of being able to concentrate on other feedback sources for monitoring purposes.

In one experiment Lee (1950b) required his subjects to read a passage first without delay and then under DAF, while maintaining accurate speech, which he defined as speech without repetitions and speech at an even rate. He suggested that this could be accomplished by cooperating

with the slowing down effect, in order to "achieve the cadence demanded by the delayed feedback" (ibid., p. 824). Under the DAF conditions, all but one of his subjects were found to have considerably increased total reading times from the no-delay condition. Furthermore, the longer the delay the longer the total time taken.

In a second experiment (Lee, 1950b) subjects were asked to repeat as rapidly as possible a single syllable (three phonemes) 40 times. Again, the total time required for the repetitions was longer under DAF.

Lee (1951) remarks that his subjects either slowed down and spoke with increased intensity, or when they attempted to maintain normal speed, they halted and repeated syllables and continuant sounds.

Black (1951), who had been investigating DAF effects independently of Lee, required his subjects to read a series of five five-syllable phrases under 11 different delays. The order of presentation of the delay conditions was rotated systematically among the subjects. Subjects were instructed to "speak naturally" (ibid., p. 57). It was found that the mean duration of each series of phrases increased progressively with an increase in delay up to a delay of 180 msec. With longer delays mean reading times decreased again, without, however, approaching the shorter zero-delay duration.

> "The effect of increasing the delay from 30 to 60 msec is proportionately greater than the effect of any other comparable 30 msec increment within the range of the study" (ibid., p. 60).

Intensity, based on the mean of 15 peak intensities per condition (three peaks per phrase), increased progressively as the delay increased up to 270 msec. However, the increases were significant only up to 90 msec.

A brief study by Rawnsley and Harris (1954) of the nature of the artificial stutter revealed, based on spectrographic analysis, that the first syllable to be repeated later, lacked the expected transition to the next syllable, whereas the final repeated syllable showed all the transitional properties of the corresponding segment under SAF.

Fairbanks (1955) required his subjects to read a passage first without delay and then at four delays presented in random order. Each subject was instructed to "read as you usually do" (ibid., p. 335) and was familiarized with the passage and with DAF prior to the experiment.

Results showed that more errors (each instance of error regardless of length or nature was tallied as one error) were made and longer total durations with and without pauses obtained in the 200-msec delay condition than in any other conditions. Vocal sound pressure (mean of graphic peaks, re the same arbitrary reference for all samples) was increased in the DAF conditions by 10 to 12 dB on the average . There was no substantial change in sound pressure

increases over the range of time delays studied. Similarly, fundamental frequency increased as a function of DAF, but changed little as a function of delay interval. The increase ammounted to "about three and one-half semitones" (ibid., p. 341).

Fairbanks termed articulatory errors and increased duration "direct effects" and mean sound pressure and fundamental frequency increases "indirect effects" (ibid., p. 341), on the basis of their different functions.

In addition, the rate of articulatory error and correct word rate was plotted as a function of delay to a relative rate ordinate, "using the mean rates for the undelayed condition as respective constants" (ibid., p. 342). If the subject increased duration yet made the same amount of errors under the various delay conditions, the rate of error decreased as a function of delay. If, on the other hand, the subject increased his errors as a function of delay while durational increases remained constant, then the rate of error increased as a function of delay. If error and duration both were to increase proportionately, the error rate function would be a flat one. Thus, the interaction of error and duration is reflected in such a measure. The actual plotted function revealed that errors increased more at 200 msec than at any other delay. The plot was based on means. Study of individual recordings revealed that

> "total disturbance in a given sample involved in many cases a seeming compromise with the interference, in which duration was 'traded' for articulation, or vice versa, but in proportions that varied from case to case" (ibid., p. 343).

The ratio of correct words per time unit decreases if either total duration is increased, or number of correct words decreased. In such a measure numerator and denominator do not tend to cancel each other. Fairbanks, therefore, suggested the following measure of DAF disturbance, however, with the objective that intercomparison of different texts could be made more objectively:

> "
> $$I = \frac{W_c D_n}{W_t D_o},$$
>
> where W_c is obtained number of correct words, D_n is normal duration, W_t is total number of words, and D_o is obtained duration, D_n/W_t being a constant for the text" (ibid., p. 344).

The experimental data obtained in the previously described study (Fairbanks, 1955) were further treated by Fairbanks and Guttman (1958) as follows: three additional sentences of the passage were analyzed as before and results described to be in agreement with the previously reported findings. Errors of articulation were located and assigned to one of the following categories: substitution, omission, addition, plus a miscellaneous category. Results showed that

> "number of instances of error varied substantially with both delay interval and type of error" (ibid., p. 21).

The number of disturbances of certain types were related to specific values of delay, i.e., the types of errors were present in different proportions for different delay intervals. It is of interest that during the 200-msec delay condition, additions were most numerous, 70% being classified as repetitions. The length of error (number of phonemes omitted, substituted, etc., in any one instance of error) was also investigated. There were more substitutions and repetitions involving two to four or more phonemes and three to ten phonemes respectively in the 400 msec delay condition than in any other condition. However, errors involving fewer phonemes occurred more frequently and were found to be most numerous in the 200-msec delay condition.

A description of the DAF literature would not be complete without mention of the following study by Chase (1958): in the first of two conditions, Chase asked his subjects to repeat the sound [b] "as quickly as possible" (ibid., p. 585) at zero delay, starting at a given signal and stopping at a signal after five seconds. In the second condition a 216-msec delay interval was introduced. Results showed that 15 of his 20 subjects repeated the sound from two to seven more times under the delay condition than under the SAF condition.

A much more detailed study of DAF effects than had been carried out up to this time was attempted by Kozhevnikov and Chistovich (1965). Five subjects were required to recite ten previously learned phrases of about 14-25 speech

sounds each twice under SAF conditions and at four different delays.

The following parameters were recorded by means of an oscillograph: contact between lips, voice, contact of the tongue with an artificial palate at three points in the alveolar and alveopalatal regions. A phoneme-per-second measure revealed that the 150-300 msec range produced maximum increases in duration; however, subjects differed greatly from each other in terms of the delay which produced maximum disturbance and in terms of the magnitude of the disturbance. For one subject almost no durational increases were found. Further analysis was therefore based on the four subjects who were affected by DAF. An examination of the oscillograms revealed

> "multiple closing and opening of one and the same contacts instead of their one-time closing and opening as in the normal speech" (ibid., p. 145)

for all values of delay. Perceptually, the result was a sequence of open syllables. Repetitions of longer speech units were also observed, the majority occurring at a delay of 470 msec. Such a repeated longer unit was reported to have been accompanied at times by multiple repetitions of shorter segments, i.e., by the multiple closing and opening described above.

Further analysis revealed that the durations of contact closure and contact opening did not change as a function of delay interval, but were approximately constant

under DAF conditions. The durations of closure and opening were approximately the same. When the number of closing-opening movements of the same contacts involved in any one incidence of repetition (provided at least two such closing-opening movements occurred in one such incidence) were counted, it was found that the number increased with an increase in delay. Total duration measures did not reflect this fact since prolongation of a sound was also found to lead to a decrease in multiple repetitions. The transition from closing to opening was found to be the only aspect of speech not delayed under DAF. More complex sequences of consonants (C) and vowels (V), if altered by DAF, were found changed as follows:

" CVC = CV + CX
CCV = CX + CV
CVCC = CV + CX + CX
where X is a reduced vowel" (ibid., p. 156).

Even if the consonant was voiceless, the opening of the closure was accompanied by voicing.

Final vowels in a series of repetitions were found to be of longer duration and to coincide with durations under normal feedback conditions. The durations of non-final vowels in a series of repetitions were found to be fairly constant and unrelated to durations of vowels in normal feedback conditions.

In each study, in which an attempt was made to place a value on the maximally disturbing delay, individual differences have been a considerable source of variance.

MacKay (1968) hypothesized that speech rate might be correlated with the delay interval producing maximum disturbance (peak shift hypothesis). In one of a series of studies he asked adult subjects to repeat sentences at a maximally fast, slow, and very slow rate of speech under two delay conditions. Number of errors (repetitions) was chosen as the dependent variable. His results did not confirm the peak shift hypothesis: there were proportionately fewer errors with diminishing rates for both delay values.

However, in another experiment in the same series of experiments when subjects repeated sentences at a maximum rate of speech with no delay and at five delay value, the peak shift hypothesis was confirmed: it was found that the subjects with the fastest rate of speech under the SAF condition exhibited the fewest repetitions under DAF. The length of delay producing maximal disruption, as calculated from a correct syllable interval (in seconds per syllable), was found to be related to the maximum rate of speech in the SAF condition.

> "The slower a subject's maximum rate of speech, the longer his critical DAF interval." (ibid., p. 819).

What little is known about local rate under DAF to this date is due primarily to the investigations of Huggins (1968). The test material in this study consisted of two sentences, one containing only short vowels, the other containing only long vowels, both preceded by a short utterance serving as a frame. The subject was familiarized

with DAF prior to the experimental readings and instructed to "speak as fluently as she could, and to try not to pause in the middle of a sentence." (ibid., p. 54). She had before her a V.U. meter and was told to maintain the same constant speaking level in both, normal and delay, conditions. The sentences were read 20 times at zero delay and then once each at 33 values of delay and finally again at zero delay.

When the total duration of each sentence production was plotted as a function of delay interval, several peaks of disturbance for both sentences were evident. The highest peaks for the sentence with the longer syllables (mean syllable duration 310 msec at zero delay) occurred at a lower value of delay than the highest peak for the shorter syllable sentence (mean syllable duration 160 msec at zero delay).

Fine analysis consisted of measuring the duration of every acoustic segment in each sentence at zero delay and at 12 values of delay from spectrograms. It was found that

> "the duration of some segments was equal to, or even shorter than, the duration with zero delay" (ibid., p. 56);

(/ɩ/ in /sɩli/). The durations of unstressed syllable segments tended to remain constant as the delay changed, while the durations of other segments grew linearly with the delay (initial and final /s/ and /ʃ/, closure of initial voiceless stops, /l/ and /r/ in intervocalic

position in a stressed syllable, long vowels in stressed syllables, short vowels in the last syllable before a terminal juncture).

> "Still other segments had a duration that did not change as the delay was changed, but was about 60 msecs longer than the duration with zero delay" (ibid., p. 56);

(occurred with all phones, sometimes in addition to the linear increase).

A total of 48 repetitions in all short sentences and 19 repetitions in all long sentences were counted, most of which occurred at 180, 240, and 360 msec. Duration of the repeated segments, which were measured and analyzed separately, correlated highly with length of delay.

An attempt was made recently by Fletcher and Yates (1971) to replicate the earlier described experiment by Chase (1958). However, Fletcher and Yates used a higher feedback level and only approximately the same delay values. Subjects had no prior experience with DAF; (it is not known if the subjects in Chase's study were naive with respect to DAF). Subjects were required to repeat each sound correctly after the experimenter and were instructed to repeat the sounds as quickly as possible but exactly as practiced. Speed was termed "important" in the instructions and accuracy "vital" (Fletcher and Yates, 1971, p. 75). The experiment was also extended to include repetitions of 19 other consonants besides [b]. Results showed that subjects repeated sounds fewer times under DAF, a contradiction of the results reported by Chase.

In order to reduce fatigue effects and to see if similar results could be obtained for vowels, two further experiments were conducted. Only 12 consonants were repeated in the first one and 12 vowels in the second one. Subjects were given practice in repeating the sounds as quickly as possible prior to the experimental condition. Again the sounds tested were found to be repeated fewer times under DAF.

A further experiment was designed in which half of the subjects were trained to repeat sounds under SAF conditions at six sounds per second (mean rate of Chase's (1958) subjects), and the other half trained to repeat sounds at three sounds per second (mean rate of subjects in the last two experiments described above). Loud clicks were used to pace the subject. It was expected that fast subjects would become faster under DAF and slow subjects slower under DAF. However, no significant differences were found between the two groups.

The researchers up to this point had not investigated the behaviour of a specific speech articulator under auditory delay. Realizing the need for such a study, Sussman and Smith (1971) designed an experiment to investigate such specific aspects of DAF behaviour. Subjects, naive with respect to DAF, were required to repeat the carrier phrase "that's a CVC month", in which the vowels /i/, /ɛ/ and /æ/ were embedded. The phrases were repeated three times each at zero delay and at four magnitudes of

delay. Jaw movements of the subjects were recorded by means of a strain-gauge transducer.

Results were as follows: extent of jaw opening for the vowels did not change significantly as a function of delay condition. However, visual inspection of the graphic records revealed

> "a slight, but rather consistent, increase in jaw lowering under the 300 msec delay for /ɛ/ and /æ/ vowel contexts" (ibid., p. 689).

Duration of jaw activity for each vowel context was significantly longer under the 100-msec condition. The longest durations at all values of delay were recorded for the /i/ vowel sentence. At a delay of 400 msec, jaw activity again approached durations recorded under zero delay. Opening and closing velocity of jaw movements was only slightly changed as a function of delay. However, though not statistically significant, jaw opening velocity for /ɛ/ and /æ/ was greatest at the 300-msec delay. For /æ/ the jaw closing velocity as well was greatest at 300 msec delay.

> "The characteristically greater closure velocity during the open vowel context was evident only for the normal and 100 msec delay conditions. Though there was a tendency for jaw closure velocities to increase as a function of vowel openness, the auditory delay altered the relative velocity relationships of opening to closing jaw movements, such that jaw opening velocities were greater under delayed feedback conditions as compared to jaw closing velocities." (ibid., p. 690).

Recently another attempt was made at relating speech rate to maximally disturbing delay. Robinson (1972)

controlled speech rate by providing external pacing. Subjects were required to count aloud flashes of lights and were

> "carefully instructed to enunciate each number clearly and evenly and to try to make the number last as long as the light stayed on" (ibid., p. 2).

Five series of 50 flashes were presented at each of eight different flash rates. Within each series the flash rate was constant, and beginning with flash number 21 of each series the subject heard his own voice delayed. Five different delays were used. A white-noise spectrum restricted to the frequency band of human speech of about 60 dB SPL accompanied the auditory feedback signal during the entire experiment in order to mask bone conduction. Analysis consisted of counting articulatory errors according to a scoring system in which each spoken number had a potential error score of three. A significant interaction between speech rate and delay was found. A longer delay produced a maximum number of errors at a slower speech rate while a shorter delay produced a maximum number of errors at a faster speech rate. The articulation score function is approximately linear.

2.3 Discussion

Obviously, not all of the results obtained by the DAF experimenters are in agreement. Chase found that [b] could be repeated faster under DAF, yet Fletcher and Yates

found their subjects were repeating single speech sounds fewer times under DAF. MacKay did not find a 'peak shift' as a result of speech rate changes, whereas Robinson did. Neither Fairbanks and Guttman nor Huggins encountered the multiple repetitions so characteristic of the subjects in the Kozhevnikov and Chistovich study. Sussman and Smith did not find that the 200-msec delay value produced a universal disruptive effect on all aspects of jaw movement as would have been predicted by the Fairbanks' maximal disturbance functions. Can these seemingly contradictory findings be explained in terms of different playback levels, subject instruction, inherent subject variability, measures employed in the analysis, or even in terms of interpretation of results? An attempt will be made to resolve some of the discrepancies in a review of some pertinent aspects of experimental design and of the underlying assumptions of the experimenters.

Amount of amplification of the delayed signal has been shown by Chase and Guilfoyle (1962), among others, to be an important variable in the degree of disturbance produced by DAF, i.e., with increasing intensity of the delayed feedback signal, increases in disturbance occur. However, as Fairbanks (1955) has pointed out, some experimenters such as Black (1951) held the intensity of the feedback signal approximately constant at the earphones by employing volume limiters, whereas others, such as Fairbanks (1955) himself, amplified the vocal output by a

specified amount by adjusting the gain of the playback amplifier. If volume limiters are used, an increase in vocal intensity would result in a higher level of undelayed feedback since the amplified delayed signal would be limited by the setting of the volume limiter. In such a system acoustic segments with a higher intrinsic intensity are possibly less subject to DAF interference than those acoustic segments with a lower intrinsic intensity. In any event, as Fairbanks (1955, p. 337) pointed out, nonlinearities of the response can be expected, though a very high setting of the volume limiter would minimize any such nonlinearities. If, on the other hand, vocal output is amplified by a specified amount and no ceiling placed on the amplified feedback signal, then vocal output intensity is approximately linearly related to feedback signal intensity. The difficulty in this system exists in finding a setting high enough not to limit vocal output by approaching pain threshold levels, or not to overmodulate the tape.

In view of these considerations can any of the above mentioned conflicting results be explained in terms of type and amount of amplification? Chase did not amplify his signal by as much as Fletcher and Yates. Does this imply that the subject with a lesser amount of delayed feedback available to him will 'stutter' rather than prolong? The question will be discussed below in terms of speech production and DAF interference mechanisms. Kozhevnikov and Chistovich kept the level of the feedback

signal constant at the earphones, but they used such a high setting that subjects can not be expected to have had any undelayed auditory feedback available to them. Kozhevnikov and Chistovich themselves explained the multiple repetitions in terms of the high level of feedback. However, Huggins' subject performed under a level of delayed feedback not much lower than that employed in the Kozhevnikov and Chistovich study but did not produce multiple repetitions. As Huggins points out, this type of explanation is inadequate. Huggins, offering an alternative explanation, suggested that perhaps speakers of Russian perform differently from speakers of English. Yet, he remarks that such a solution is even less likely than the solution based on the different intensities of the feedback signal. In view of the lack of data, however, such a hypothesis remains to be tested. Huggins arrived at the conclusion that the most likely cause of multiple repetitions, both in the Chase, and Kozhevnikov and Chistovich studies is the set adopted by the subjects, i.e., does the subject know the purpose of the experiment? Huggins appears to have arrived at this conclusion because he himself could produce such multiple repetitions at will, yet had not encountered them in his subjects. Huggins seems to be suggesting that subjects usually resist producing multiple repetitions. This hypothesis will be debated further in a discussion of the underlying assumptions of the experimenters.

Before proceeding with this discussion, one further

possible explanation for the frequent multiple repetitions in the Kozhevnikov and Chistovich study suggests itself. Although it has not been considered by the experimenters mentioned thus far, the artificial palate could have been responsible in that it eliminated a portion of the tactile feedback, thus reducing the undelayed neural feedback available for control and causing what Fairbanks termed 'uncontrolled oscillations'.

Lee conceived the speech mechanism as being

> "somewhat like a machine gun, repeating aurally monitored units as long as the trigger is held down" (Lee, 1951, p. 54).

He suggested that the aurally monitored unit probably is the syllable, remarking that DAF-induced phoneme repetitions were not observed. However, phoneme-by-phoneme monitoring was hypothesized by Lee to take place via kinesthetic and tactile channels.

Black similarly considered

> "that the duration of the syllable may have an important bearing upon the effects of delayed side-tone" (Black, 1951, p. 58).

The hypothesis was based on the fact that the delay of 180 msec (approximately the duration of the syllables produced by Black's subjects under DAF) resulted in the longest reading time. Due to the disproportionate increase in reading time with an increase in delay from 30 to 60 msec, Black did not exclude the possibility that the phoneme may be a monitored unit as well.

Both authors arrived at similar conclusions via different observations. Lee's hypotheses are based on the nature of the disturbance itself, whereas Black arrived at his conclusions by relating selective delay intervals to magnitude of disturbance.

Basic to these hypotheses is a unit-by-unit monitoring system, an assumption reiterated by Chase: each produced speech unit "feeds back to the speaker and contributes to the release of the next speech unit"(Chase,1958, p. 584). DAF affects the monitoring system such that each speech unit tends to circulate one or more times due to the temporal discrepancies of vocal output and feedback signal and "changes in the physical aspects of the speech signals themselves" (ibid., p. 588).

Fairbanks, commenting on the idea of a possible relationship between the maximally disturbing delay interval and the duration of conventional speech output units such as the syllable, remarks that the data are meager, and that "the criterion of disturbance ... is not yet established" (Fairbanks, 1955, p. 342). Influenced by cybernetic theory, he conceived of the speaking system as a closed cycle system:

> "A closed cycle system, or servosystem ... employs feedback of the output to the place of control, comparison of the output to the input, and such manipulation of the output-producing device as will cause the output to have the same functional form as the input" (Fairbanks, 1954, p. 135).

One part of the system is a comparator which receives the feedback signals as well as the input. Here the error signals are calculated and sent to an effector unit via a mixer. This mixer determines the effective driving signal, which is then sent to the effector to modify the process of speech production until the error signal is zero. The error signal may travel several times around the comparator-mixer-effector loop being progressively reduced until it equals zero. An important part of the system is a storage component which receives input from the comparator. It is able to calculate the time at which the error will reach zero. The next unit of input can therefore be started before the error signal reaches zero, so that the effector does not have to wait for a new driving signal.

Fairbanks suggested that the unit of control, referring to the unit of speech production control, should not be identified with any conventional linguistic units.

> "It seems more likely that the unit of control, although related to such units, is more fundamental, and that it may consist, at any given time, of one or a number of such conventional units, or of a fraction of the smallest phonetic unit." (Fairbanks, 1955, p. 342).

He referred to the unit as a

> "semiperiodic, relatively long, articulatory cycle, with a correlated cycle of output." (Fairbanks, 1954, p. 138).

Under DAF at any given time delay, the number of control units which have periods critically related to the time delay might be inferred from the DAF functions.

Consistent with the above model in which each succeeding segment depends on the feedback from the previous one, Fairbanks and Guttman (1958) present their view of the DAF interference mechanism. Pertinent to their model is the fact that predominatly double repetitions occurred in their data. Thus, they only account for double repetitions and consider that the occasional multiple repetitions are

> "uncommon ... person-linked, and give the impression of wild and uncontrolled oscillation of the vocal mechanism" (ibid., p. 20).

Fairbanks and Guttman suggest that the primary form of DAF disturbance are repetitions, which are described as "automatic responses to misinformation in the feedback complex" (ibid., p. 19). A repetition automatically results when the subject produced segments A and B and heard A during B, thus triggering another B and temporarily restoring normal feedback relationships. If the system were not aperiodic, subjects would be expected to produce A B B C C D D etc. under DAF conditions.

A similar mechanism of DAF interference is proposed by Huggins, which does, however, try to account for multiple repetitions. Huggins suggests that DAF causes the control mechanism to lose its 'place' in the program, in which case "the target articulation of the previous instruction is continued until the place is re-found" (Huggins, 1968, p. 59). If the control finds its place

from undelayed sources of feedback (either neural or bone-conducted), then only a slight lengthening occurs. If, on the other hand, the place is found from the delayed auditory feedback source, "then an extra segment may be inserted into the articulatory sequence, with duration equal to the delay" (ibid., p. 59). A necessary condition for such insertion is probably the similarity of the produced and monitored delayed signal, which can only occur in a rapid sequence of articulations. If the place is re-found during a slower sequence of articulation from the delayed source, drawling will occur. Huggins has not yet published the promised companion paper in which he would try to discover "which articulatory sequences generate most repetitions" (ibid., 1968, p. 59). Thus, the above hypothesis remains untested to date.

Kozhevnikov and Chistovich viewed the DAF-induced repetitions as well as the prolongations, regardless of the conditions which determine which one of these DAF-induced disturbances will occur, as basically the same phenomenon. They disagreed with the conclusion of Fairbanks and Guttman that speech production units do not coincide with any known linguistic units. It is this disagreement which prompted Kozhevnikov and Chistovich to carry out their own DAF investigations. As has been stated previously, Kozhevnikov and Chistovich did not find multiple repetitions uncommon, in fact they found them to be the most clearly expressed effect of DAF in their data.

Clearly, these repetitions had to be accounted for in an explanation of DAF interference.

Kozhevnikov and Chistovich, like Fairbanks and Guttman, found repetitions of longer units occuring more frequently at longer delays. Did the fact that the repeated fragments were extremely varied necessitate the conclusion that speech units should not be identified with any of the conventional linguistic units? Kozhevnikov and Chistovich argue that because of the variability of these fragments nothing can be known from these repetitions. Furthermore, multiple repetitions were at times superimposed on repetitions of longer fragments, which caused the above authors to consider that the effects may be of a different nature.

Rejecting Fairbanks' assumptions, Kozhevnikov and Chistovich proposed the following model: through the automatic activity of the nervous centers which control the antagonistic muscles of the lips, tongue, and vocal chords a CV-syllable is formed. Delaying the auditory feedback excludes temporarily the controlling influences of the higher nervous centers. In this model the period of fluctuating closing-opening movements as a result of DAF would be determined by the activity of the nervous centers themselves and not by the delay interval. Such a model was derived by the above authors from a number of observations: the frequently observed lack of coordination between vocal fold activity and articulation was considered evidence for

DAF-induced fluctuations in parallelly operating centers of antagonistic muscles. That DAF occurs at a reflex level was deduced from the transformation of the CCV structures into CV structures (on the basis of their previous research these same authors hypothesized, CCV structures constitute one cohesive unit in the higher centers). Thus, it became necessary to consider that CV structures are basic while more complex combinations such as CCV are made up of CV complexes

> "organized so that the following complex begins before the preceding one is able to finish." (Kozhevnikov and Chistovich, 1965, p. 158).

Further confirmation of the open syllable as a cohesive unit was provided by the fact that no slowing was observed from C to V but only from V to C. Sussman and Smith's finding that jaw opening velocity was at times even facilitated by delay, whereas velocity of jaw closing was slowed, seems to further substantiate the above hypothesis; (Sussman and Smith have not related their findings to the Kozhevnikov and Chistovich model). It must be considered a possibility, though, that the Sussman and Smith results on English are not strictly comparable with the results obtained on Russian. However, if Kozhevnikov and Chistovich are correct in their assumption that DAF operates at a reflex level and that the opening-closing movements are basic to speech production and that they can be studied under DAF in the temporary absence of higher control, then the phonological system of the speaker would not be a

crucial aspect.

Accepting the CV-syllable as a cohesive structure, it follows that repetition of the same open syllable constitutes preservation of the same structure. Since the time of existence of a structure grew with an increase in delay, it was

> "necessary to consider that the essence of the effect of the delay of an acoustical signal consists not of the calling of repeated movements but in the delay of the transition from one structure of movements to another. Actually, in the case of a delay of the acoustical signal man can arbitrarily either stutter or draw out the speech. If he hurries, stuttering occurs; if he tries to speak slowly, the speech is drawn out. <u>However, regardless of what he does, the change of structures is slowed and the time occupied by the pronunciation of the phrase increases.</u>" (this author's emphasis) (Kozhevnikov and Chistovich, 1965, p. 155).

Crucial to the theory proposed by Kozhevnikov and Chistovich is the fact that time of existence of a structure (number of repetitions in a series) was related to the delay interval. Huggins states that durations of repeated segments in multiple repetitions in the Kozhevnikov and Chistovich study were not related to the delay interval, a finding which he considered along with the multiple repetitions not in agreement with his findings. In view of the above advanced hypothesis, that multiple repetitions are one expression of prolongation, it is crucial that the length of the multiple repetitions, and not each segment within the multiple repetitions be related to the delay

interval.

The seeming contradiction of the Chase, and Fletcher and Yates studies might now be considered. In the former study, subjects exhibited multiple repetitions; in the latter they slightly prolonged. Huggins' model would predict that the behaviour of Chase's subjects was governed by the delayed source of feedback, whereas the subjects in the Fletcher and Yates study relied on undelayed feedback. However, Fletcher and Yates, as has been mentioned, amplified the feedback signal more than Chase, thus reducing possible undelayed feedback. Huggins' model of undelayed feedback also includes neural feedback. We would have to assume that the Fletcher and Yates subjects relied entirely on neural feedback. Does this imply an inherent difference in the two sets of subjects, or does it imply that a higher level of delayed feedback causes subjects to rely more on neural feedback? Kozhevnikov and Chistovich used a very high level of feedback, yet their subjects performed like the Chase subjects. In view of the unusually high number of subjects in the Fletcher and Yates study in all three experiments, inherent between-subject differences does not appear to be a reasonable explanation either. Huggins, as mentioned previously, believed that the set a subject adopts is responsible for multiple repetitions. He mentioned subject knowledge of the purpose of the study as one possible explanation. It cannot be disputed that instruction to the subject influences his attitude. Chase

stressed accuracy and speed. Is it possible, as Huggins appears to be implying, that, normally, subjects resist multiple repetitions unless they believe that this is the desired behaviour in the study? It is reasonable to assume that multiple repetitions consitute a behaviour without control. Fairbanks had already termed multiple repetitions 'uncontrolled oscillations' and Kozhevnikov and Chistovich made more explicit what is meant by 'uncontrolled'. It follows, since each system normally strives for control, that subjects could indeed by expected to fight a loss of control. As Lee (1950a) already pointed out, more than two minutes of exposure to DAF is physically tiring. A subject then has to be either unable to maintain control or willing to lose control.

Despite the previous discussion, we are not necessarily confronted with contradictory results. If one accepts the hypothesis of Kozhevnikov and Chistovich, multiple repetition of the same structure constitutes slowing of the structure and prolongation of the same structure also constitutes slowing. Because the required task in the Chase, and in the Fletcher and Yates experiments consisted of repeating <u>one</u> structure, repetition of this <u>one</u> structure, either fast or slow, constitutes slowing. Since a transition to a new structure did not need to be accomplished, nothing can be known from the results. Let us consider the case where the subjects would be asked to repeat two structures such as [b] and [d] in rapid alternation,

it being understood that an epenthetic vowel could be part of each structure, as it is difficult to say the speech sound without such an insertion. The hypothesis of Kozhevnikov and Chistovich would predict the following results: the subjects in the Chase study would have produced rapidly [bbbbdddd] (any number, depending on the delay), but not [bdbdbd]; whereas the subjects in the Fletcher and Yates study would have produced [bˀbˀb dˀdˀd bˀbˀb dˀdˀd] where an epenthetic vowel (at the release of the closure) could also be prolonged. Under a series of delays the relative number of repetitions of a structure produced by the former set of subjects would be related to the relative amount of prolongation exhibited by the latter set of subjects. Only if the Chase subjects would produce [bdbd] in rapid succession could the results of the studies be considered contradictory.

Why some subjects predominantly exhibit one type of slowing behaviour and other subjects the other, or what determines the occurrence of both behaviours in one subject at different times must remain largely speculation until more is known about these behaviours themselves, until their interaction can be more precisely described and until we know more about the relation of these behaviours to normal rate of speech and about rate of speech itself.

Robinson conceived timing in speech as being dependent on the speech itself as opposed to a biologically determined rhythm. He interpreted his findings as evidence

for such a model. As pointed out earlier, MacKay did not find a peak shift for disturbance with change in rate under DAF. However, since MacKay only counted repetitions as errors and Robinson counted all errors, including omissions, their results are not strictly comparable. Robinson also tallied "striking quavers within a vowel" as error (Robinson, 1972, p. 2). The difficulty with such a scoring system is to decide when a quaver is striking and when it is not. As was already foreseen by Chase, although he was specifically speaking about repetitions of phoneme-sized units, a graphic speech recording might be necessary to reveal if an actual repetition took place (Chase, 1958, p. 589). Thus, many actual prolongations are possibly included in Robinson's error count and many actual repetitions might have been missed. One must therefore view Robinson's linear articulation score function with some reservations.

MacKay, having discovered that subjects with a slower maximum rate of speech showed a greater tendency to stutter under DAF than subjects with a faster rate of speech, whereas subjects voluntarily decreasing their rate under DAF showed a decreased number of repetitions with decreasing rate, suggested different mechanisms for the normal speech rate under SAF and prolongation under DAF. Altering rate under DAF must then be viewed not as a true rate change but as a trading of one type of slowing for another, keeping in mind that the previously established framework in which repetitions are seen as slowing of the

transition from one structure to the next. Thus, with an increase in one slowing behaviour (prolongation) a decrease occurred in the other (repetition). It appears that the two forms of slowing are mutually exclusive behaviours at any one time. Why subjects with a slow maximum speech rate in the SAF condition should exhibit more of the repetitive type slowing cannot be explained on the basis of the data available.

In view of the above considerations one must seriously reconsider Fairbanks' (1955) suggestion that articulatory error (substitution, omission, repetition) is the primary form of DAF disturbance. In order to assess the interaction of the two types of slowing behaviour, a fine analysis of local rate is needed. In addition, global rate must be varied under SAF as well as under DAF and the effects on local rate assessed. If slower rate under DAF can be shown to exhibit different slowing patterns from slowed rate under SAF then MacKay's hypothesis postulating different mechanisms for rate changes under SAF as opposed to DAF can be substantiated.

Allen (1973) has pointed out that "although speech rate and phonological length are the major determinants of most segmental duration" (ibid., p. 222), there are neuromuscular and biochemical constraints that are not entirely under the control of the higher nervous centers. The interaction between phonological length and phonetic duration is complex. Allen's diagrammatic speech production

model contains a structured phonological input into an output buffer. The output buffer provides the input for a speech motor control program (transfer function) which determines partly the articulator movements. Allen suggests that when we measure the output of the system, the variability due to segmental timing control error in the motor control program (speech rate variation) can be isolated from peripheral variability (neuromuscular and biochemical constraints and measurement error). He proposes a statistical model to separate these variance components. However, this model can be applied only if the speaker attempts to keep his rate fixed in the repeated production of the same utterance.

How does DAF fit into Allen's speech production model? Does DAF operate on the motor control program? Since the segmental timing control takes place at this level, voluntarily slowed speech (the segmental timing clock is voluntarily set to a slower rate) and DAF-slowed speech (DAF determines the rate of the clock) would exhibit similar patterns, although perhaps greater variability can be expected under DAF if it is assumed that the clock is slowed whenever it comes under DAF control and speeds up when it escapes the DAF influence (i.e., after a pause). Allen predicts specifically a greater variability of segment durations when a local decrease in clock rate takes place (during stressed segments). Although he refers to repeated utterance of the same phrase keeping the rate

constant, it should also be true in a comparison of stressed segments during voluntarily slowed speech if it is assumed that, in both cases, slowing of the segmental timing clock takes place, provided the rate of the clock does not differ too much from one condition to the other. However, one cannot be certain, even if such a clock exists, that DAF operates on segmental timing control at the motor control programme level such that the clock itself is slowed and linear increases in duration of segments result, provided DAF is effective all the time. It has already been mentioned that DAF does not affect the speech production to the same extent from one second to the next but varies with undelayed feedback available to the speaker at a particular moment. Furthermore, it is known that certain types of sounds (continuants) are more affected. This fact cannot be explained by a slowing of the neural clock. If DAF operates selectively on manner of articulation, it is reasonable to assume that DAF interferes with the execution of the commands at a level below that of the motor control program, i.e. the reflex level proposed by Kozhevnikov and Chistovich (1965).

Perhaps DAF is active only at this level or in addition interferes with segmental timing control. Study of segment durations under voluntarily altered rate and under DAF should provide some insights into these questions.

Chapter 3

AIMS OF THE INVESTIGATION

In general, this research study is designed to investigate the durations of acoustic speech segments at a speaker's normal rate of speech, a voluntarily slowed rate of speech and under DAF.

As has been shown (Chapter 2), measures of DAF disturbance have been for the most part of a very global nature in the past. A better understanding of DAF-induced speech disturbances based on a more detailed analysis could lead not only to new insights into speech production and the role of auditory feedback, but also possibly to an understanding of stuttering as a pathological phenomenon. A review of the available DAF information as it relates to stuttering can be found in van Riper (1972, pp. 382-403). Although Huggins (1968) carried out such a detailed analysis, it remains to be tested if the patterns his subjects displayed under DAF are universally true, and furthermore, it is not clear whether these patterns are typical of a slower rate of speech under SAF as well as under DAF. In other terms, if they differ from voluntarily slowed speech, how do they differ?

The experiment reported here is designed to isolate those aspects of the temporal patterning of speech which are specific to DAF.

Specifically, the following hypotheses will be tested:

(a) When speech is slowed voluntarily under SAF, not all sounds are prolonged by the same amount. The pilot study reported below indicates that in comparison with speech at a normal rate, vowels are prolonged more than consonants. Within the class of vowels, close vowels are prolonged more than open vowels. Within the class of consonants, continuants are prolonged more than obstruents.

(b) When speech is slowed as a result of DAF, not all sounds are prolonged by the same amount. The pilot study indicates that open vowels and fricatives are prolonged the most and that position within the syllable and within the sentence selectively affects the durational increases of each sound class.

(c) When voluntarily slowed speech under SAF is compared with speech under DAF, it is evident, as indicated by the pilot study, that slowed speech under SAF follows an orderly pattern in which stress, degree of vowel openness and possibly vowel length determine durational increases. On the other hand, durational increases due to DAF are determined by articulatory parameters such as voicing, closure, position in the articulated sequence as well as degree of vowel openness.

Chapter 4

EXPERIMENTAL APPARATUS AND PROCEDURES

4.1 Subjects

One female subject served in the pilot study and five female and five male subjects served in the main study; their ages ranged from 21 to 32 years. All subjects were university undergraduates or graduates in the fields of Linguistics, Audiology and Speech Science, or Medicine. Five male and two female subjects in the main study were naive with respect to DAF. The pilot study subject and the three remaining female subjects had been briefly exposed to DAF in the past.

The two requirements for the selection as subjects were evidence of bone- and air-conduction thresholds within 25 dB of 0 dB as defined by the ISO Standard of 1964 for pure tones of frequencies 250-4000 Hz, and notable speech disturbance under DAF.

4.2 Experimental Design

The pilot subject read a short connected passage (pilot study passage, see Appendix B) under SAF at a normal rate and then at a slow rate. Following those two readings, the subject read the same passage under DAF at 200-, 180-, 220-, and again at 200-msec delay. The delay producing the longest reading time was chosen for analysis. For the SAF condition the subject was instructed to read the passage

at her normal rate and then to read it slowly. No inctructions were given for the DAF conditions.

For the main study, the hearing of all subjects was screened at the Audiovestibular Unit of the Vancouver General Hospital to establish if subjects' thresholds fell within the limits specified above.

The subjects first practised reading under DAF to exclude novelty effects. Subjects then read a short passage (screening passage, see Appendix C) under SAF. If the subjects made any articulatory errors, they were instructed to read the passage again. Subjects then read the passage under DAF over a range of delays. The range for the female subjects was 200-310 msec; the range for the male subjects was 150-240 msec. The sound pressure level at the earphones varied with the vocal output of the subjects. The same playback setting was used throughout the experiment. Subjects were instructed not to change strategies while reading, i.e., they were not to alternately slow down and speed up but to try to read as naturally as possible.

The total reading time of each subject was determined for the SAF and DAF conditions, timing the duration of the recorded passage to the nearest half a second with a stop watch. In addition, articulatory errors were counted. Each instance of error, regardless of length or type, was tabulated as one error. A second observer carried out an independent articulatory error count on half the subjects for the purpose of checking reliability of measure-

ment. In their count the two observers did not differ by more than one error on any of the individual error scores. The two measures and the following ratio were plotted for each subject as a function of delay:

$$\text{Ratio} = \frac{\dfrac{\text{DAF Total Duration}}{\text{Correct Words}}}{\dfrac{\text{SAF Total Duration}}{\text{Total Words}}}$$

The delay at which the highest ratio occurred for a particular subject was chosen for that subject as delay value to be employed in the main experiment.

Subjects returned on a subsequent day for the main experiment. They were required to read a short connected passage (main experiment passage, see Appendix D) in the following four conditions:

(a) (Condition I)
under SAF at a normal rate (NORMAL);

(b) (Condition II)
under SAF at a voluntarily slowed rate (SLOW SAF);

(c) (Condition III)
under DAF (at the selected delay for that subject) at a slow rate, such that the DAF paces the subject (SLOW DAF);

(d) (Condition IV)
under DAF (at the same delay as above) at a maximum rate (FAST DAF).

The SAF conditions always preceded the DAF conditions in the above order. The order of the DAF conditions was reversed for every second subject.

Following the four above tasks, the entire procedure was repeated, but this time the subject was asked to read four sentences, each embedded in an identical frame (main experiment sentences, see Appendix E). One sentence contained predominantly open back vowels (/ɔ/ sentence), one contained predominantly close front vowels (/i/ sentence), one contained predominantly open front vowels (/a/ sentence), and one contained predominantly close back vowels (/u/ sentence). The four sentences were presented in random order.

Subject instructions were minimal and informal. Subjects were told, for the SLOW condition, not to slow their speech by producing longer pauses between words, but to prolong the actual utterances and to try to maintain the same speed throughout. Subjects were then asked to practice this. If the subject was performing in the required manner, no further instructions were given for this task. If a subject did not perform the task as instructed and as practised, he was asked to repeat the task. For the SLOW DAF condition the subject was given a demonstration by the experimenter of what was meant by: "Let the DAF pace you" or "catch up with you". For the FAST DAF condition subjects were told to read as rapidly as possible and to try not to omit any words.

4.3 Instrumentation and Arrangement

The system employed in the production of the synchronous and delayed feedback consisted of a Uher microphone Model M516, a Körting Constellation 88 Model MT 3644 1/4-track tape recorder with movable playback heads, and earphones designed for DAF experimentation, namely such that airconducted synchronous feedback is maximally occluded. The vocal output was recorded simultaneously on a second tape recorder (Scully Model 280-2) via an Altec 681 A LO microphone.

The subject was seated four to six inches from the microphones in an IAC 1204 soundproof room. The tape recorders were outside the soundproof room. The operator was able to observe the subjects through a double-glass window and monitored the subject's speech continuously.

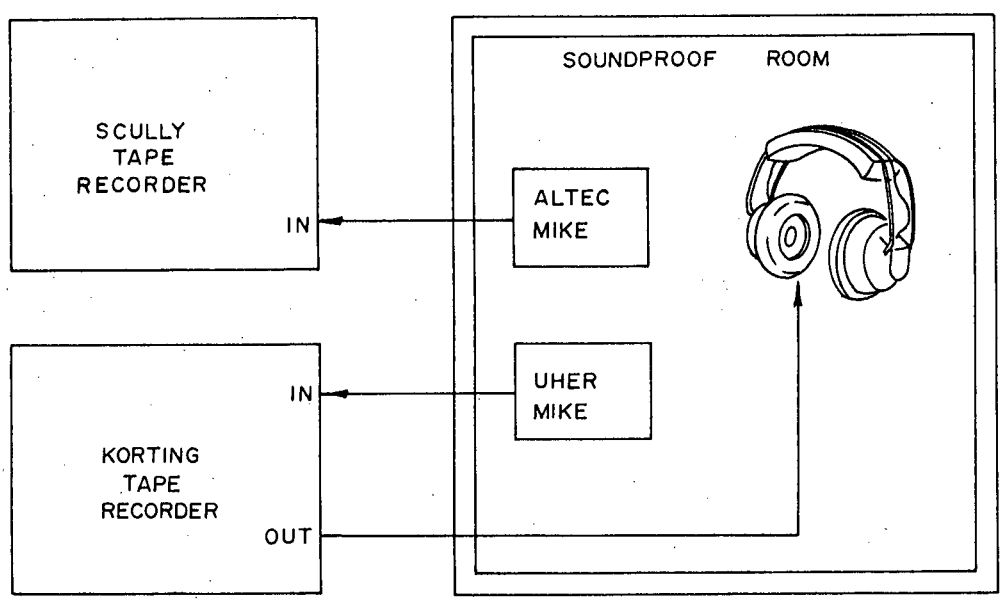


Figure 4.1. Block diagram of instrumentation for producing the delayed feedback and for recording.

4.4 Calibration

Recordings were made on both recorders at a tape speed of 7.5 inches per second on Ampex 631 professional tape. The recording level on the Scully tape recorder was approximately 0 dB as registered on the VU meter of the recorder. The playback level of the Körting tape recorder at the setting chosen was 108 dB SPL at the earphones when a 1-kHz pure tone recorded at 0 VU was played back on the Körting tape recorder. When white noise recorded at 0 VU was played back on the Körting tape recorder at the same setting, the playback level at the earphones was 95 dB SPL. Sound pressure level was measured on a Brüel and Kjaer Type 2203 Precision Sound Level Meter and Brüel and Kjaer Artificial Ear (Type 4152) with a 6 cm^3 coupler.

The precise delays on the Körting tape recorder were established as a function of the playback head position. A millimeter scale was fastened along the path of the playback head. The earphones were placed close to the microphone. A sharp click produced at the microphone was recorded at 7.5 inches per second. The microphone picked up the click repeatedly from the earphones until the click was damped. The procedure was repeated for ten different positions of the playback head. Mingograms of the initial click and its echos were then made for each setting and the distance from click onset to fourth echo onset measured in msec. The distance was divided by four. The delay

durations so obtained were plotted as a function of play-back head position (in millimeters). The delay function was approximately linear and distance values along the scale could be converted to delay values from this graph.

Chapter 5

ANALYSIS OF DATA

5.1 Instrumentation

Mingograms were made of the tape recorded speech, displaying the following four signals:

Channel 1 speech wave signal

2 duplex oscillogram

3 fundamental frequency of speech wave

4 log of average speech power

Figure 5.1 shows a schematic of the set up used to produce the mingograms.

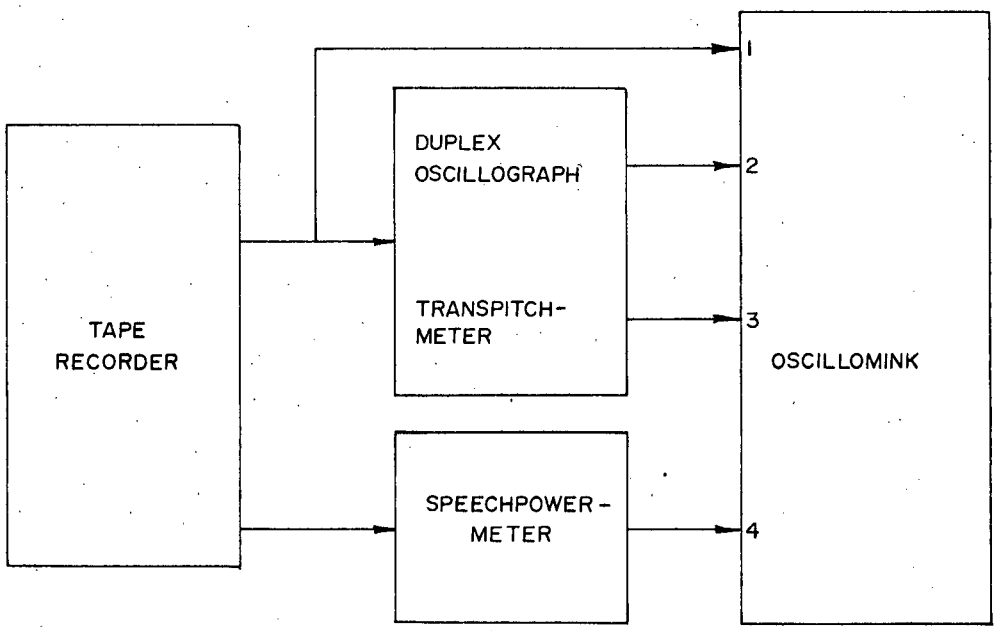


Figure 5.1. Block diagram of instrumentation used to produce mingograms.

The tape recorder used in the above setup was a REVOX Model A77. The input to Channels 2 and 3 was provided by a Frøkjaer-Jensen transpitchmeter. The intensity or speech power circuit (Channel 4) was similar to the one developed by Peterson and McKinney (1961).

Speech utterances were then segmented utilizing the above four traces on the mingograms. The duplex oscillogram yielded the most effective basis for segmentation because the higher frequencies are converted to a negative-going intensity curve. In addition, spectrograms were made on a Kay Sonagraph and a scale factor calculated to convert acoustic segment duration from the spectrograms into msec. The transcription system used in this study is described in Appendix F.

5.2 Segmentation Criteria

Vowels: the boundaries of vowels could be fairly accurately determined by comparison of all four traces (see Figure 5.2). Some difficulties were encountered if the vowel was preceded or followed by /l/, /r/, /w/, /j/, /ð/, or /v/. In these cases, spectrograms were made and the onset and ending of formants used as segmentation cues. The criteria will be discussed below for the individual phones involved in the ambiguities.

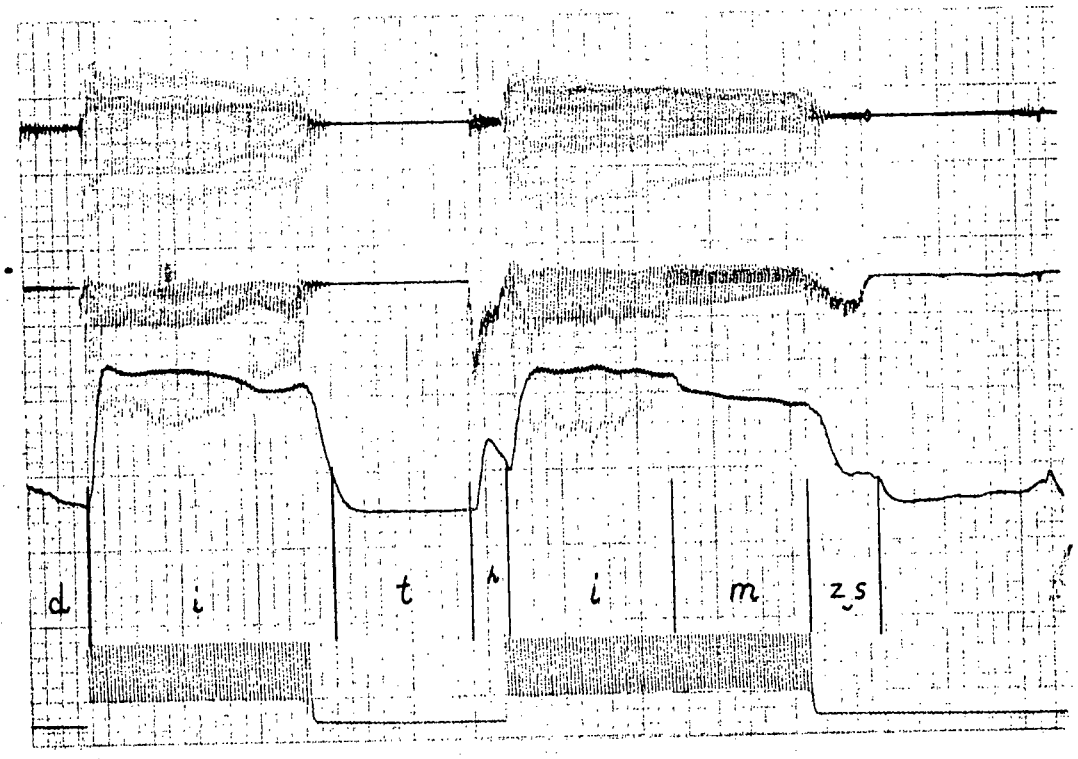


Figure 5.2. Mingogram of part of the phrase /gridi timz/.

V/l/: the momentary decrease in energy between a vowel and the following /l/ reflected by a dip in the average speech power trace, as well as changes in formant patterns on spectrograms, were used to determine the onset of /l/. If the pronunciation was as follows: /fuəl/, the /ə/ was considered as part of the /l/, because /ə/ was not part of the other occurences of /u/, whereas in some instances of /l/ a /ə/ preceded the /l/, i.e., in /pipl̩/.

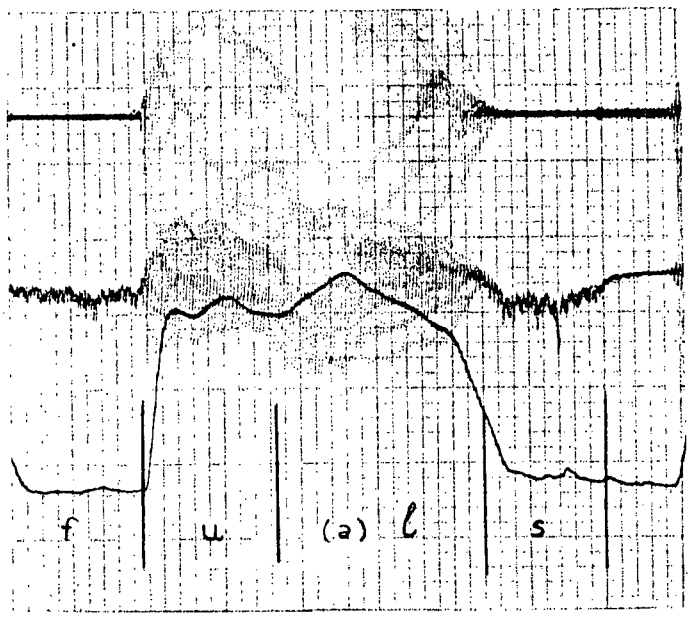


Figure 5.3. Mingogram of the phrase /fuls/.

V/r/ and /ɚ/: Segmentation of /r/ after a vowel could usually be made on the basis of spectrograms. Transition was considered part of the vowel, making segmentation more reliable. /ɚ/ occurred only after a consonant and was thus easily segmented on the basis of abrupt changes on the mingograms.

Initial /j/, /l/, /r/, /w/: The ending of these phones could be consistently determined by a change in the duplex oscillogram trace as illustrated in Figure 5.4.

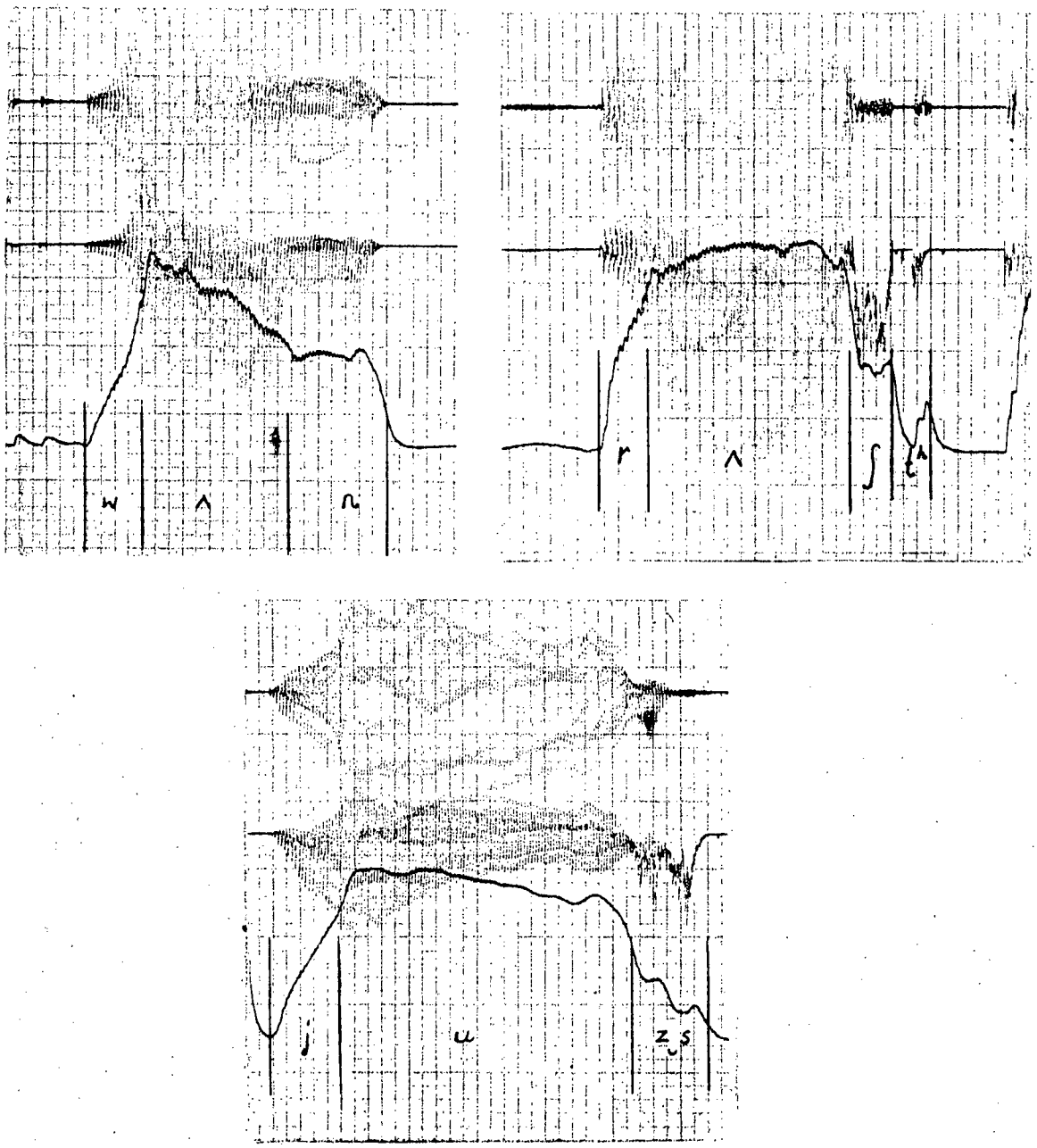


Figure 5.4. Mingograms of the phrases /wʌɹ/, /rʌʃt/, and /juz/.

/ð/ and /v/: Both a change in the duplex oscillogram trace and absence of vowel formants on spectrograms were used as segmentation criteria (Figure 5.5).

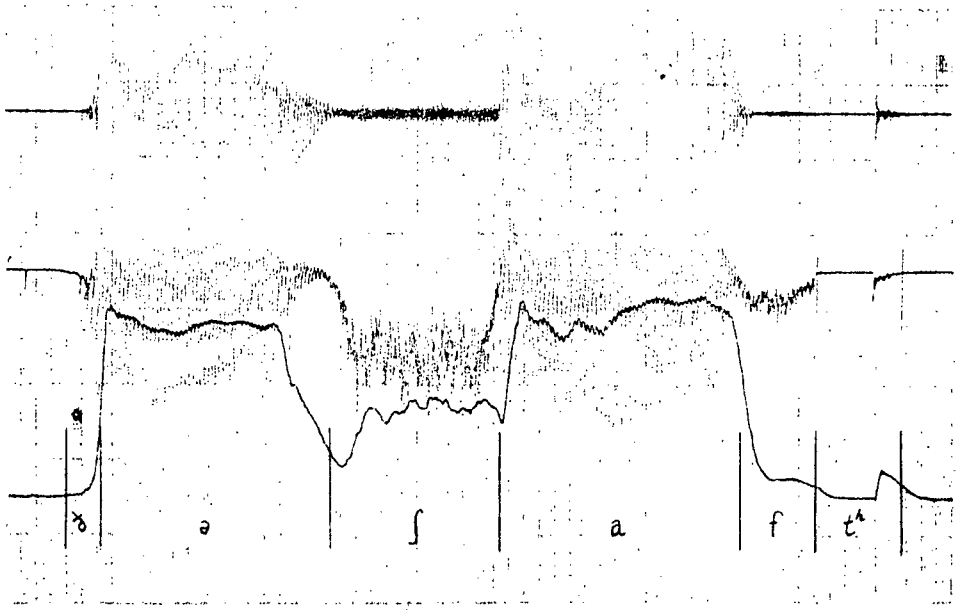


Figure 5.5. Mingogram of the phrase / ðə ʃ aft/.

Nasals: Boundaries of nasals were determined by the onset and ending of the regular pattern which nasals display on the speech wave trace and duplex oscillogram trace. In case of doubt, spectrograms were made, on which the formants of nasals could be easily discerned (see Figures 5.2, and 5.4).

Stop consonants: Beginning and end of plosion could be determined fairly accurately from the duplex oscillogram trace (negative-going intensity curve) in conjunction with the speech wave trace. However, determination of beginning of closure for initial stop consonants

proved extremely difficult in some cases where no energy registered at the onset of the closure and a pause or inhalation preceded the closure. Arbitrary decisions had to be made in some cases (see Figure 5.2, 5.4 and 5.5).

The fricatives: /f,s,z,ʃ,ʒ,θ,h/: the boundaries of these fricatives were defined as the beginning and ending of the negative-going intensity curve on the duplex oscillogram trace (see Figures 5.2, 5.3, 5.4 and 5.5).

5.3 Segment Rejection Criteria

When segments were added to the expected sequence of articulation, these segments were not used in a comparison of segment durations across conditions; i.e., in the comparative analysis described in Chapter 6. Such added segments were sometimes the result of a correction as in the following example, where an incorrectly produced word (underlined sequence) was repeated correctly: /si ðə <u>triə</u> θri/, part of the sequence "... see the three greedy teams ...". Sometimes extra segments were added, not part of the text and unrelated to the text (underlined segments): /ɒsg<u>ə</u>m<u>v</u>hɒks /, the sequence "... awesome hawk's ...".

If the added segment was the result of repetition under DAF, the following criteria were used to select one occurrence of the segment for purposes of the comparative analysis and reject the other occurrence or occurrences: when one repetition in a series of repetitions was correctly articulated, generally with a non-reduced vowel, even though

perhaps drawled, and the other repetitions were reduced or incorrectly articulated versions of the intended sequence of articulation, either in terms of vowel quality or voicing error or some other linguistic feature error, the correctly articulated intended articulatory sequence was selected for the comparative analysis and the rest of the repetition or misarticulation (underlined segments in the following examples) discarded: /gridɩdi/ (greedy), / pábasàsəsəs / (passes), / gəkəkəkaházəsəs| / (castle).

If the repeated sequences of segments in one instance of repetition were articulated correctly in each repeated sequence, the sequence with the highest speech power trace was selected for the comparative analysis (see Figure 5.6).

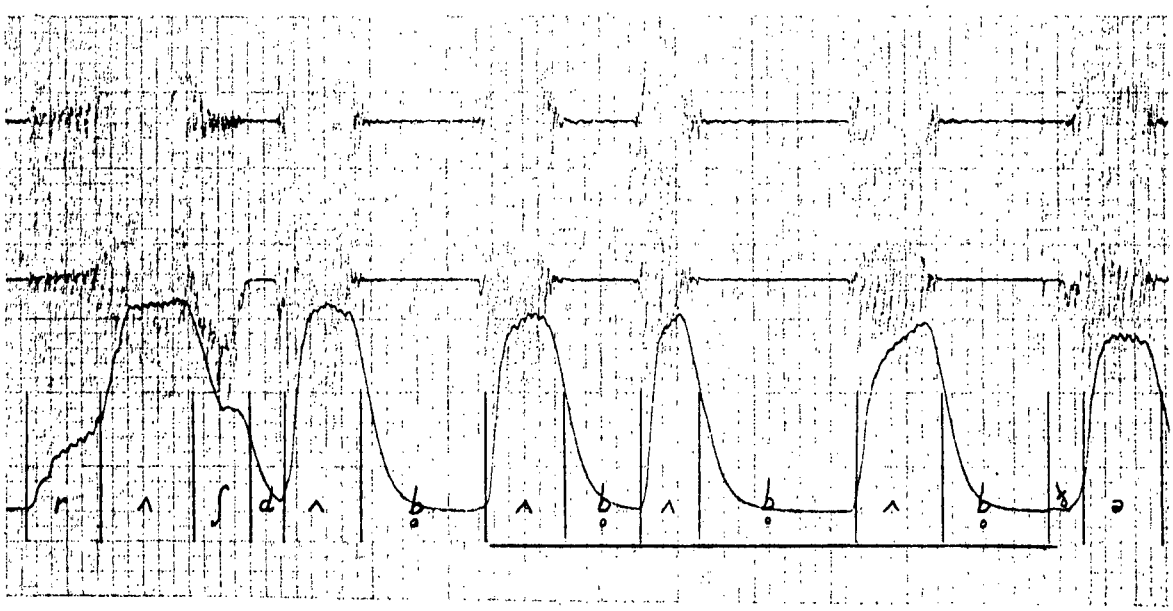


Figure 5.6. Mingogram of the sequence /rʌʃdʌb̥ʌb̥ʌb̥ʌb̥ ðə /.

If only one acoustic segment was repeated both times correctly the last occurrence of the segment was chosen for the comparative analysis.

The maximum error which may result from the above rejection criteria was estimated for the two subjects, who repeated most frequently under DAF, as follows: the sum of the longest alternative acoustic segments was divided by the sum of the acoustic segments which were not involved in ambiguities plus the sum of the shortest alternative acoustic segments. In the worst case (subject 1) the error factor amounted to 10%.

Chapter 6

RESULTS AND DISCUSSION

The results of this study will be reported and discussed simultaneously in this chapter. This mode of presentation was chosen because of the complex nature of the results. The pilot study, the screening study, the main experiment, and the summary and conclusions are presented in 6.1, 6.2, 6.3, and 6.4 respectively.

6.1 Pilot Study

The mingograms of the recorded pilot study passage were segmented in the manner described in Chapter 5. Ten phrases, beginning with the second phrase in the passage (Appendix B) were analyzed in the NORMAL, SLOW SAF and DAF (200 msec only) conditions.

Analysis of the absolute durations of the acoustic segments revealed that, compared to the SLOW SAF condition,

(1) vowels in stressed syllables, except for the open vowels /æ/ and /ɔ/, were of shorter absolute duration in the DAF condition, the greatest difference being observable in the close front vowels;

(2) open vowels were of longer absolute duration in the DAF condition, regardless of stress;

(3) fricatives and voiced stops were of longer absolute duration in the DAF condition, the greatest difference occurring in initial position;

(4) voiceless stops were of longer absolute duration in the DAF condition, the greatest difference occurring in final position.

'Stressed syllables' were defined in the pilot study as those syllables expected to carry primary stress in the NORMAL SAF condition, according to English phonological rules. Because stress could not be defined in terms of English phonological rules in the SLOW SAF and DAF conditions, it was defined only with reference to the NORMAL SAF condition, i.e., a vowel in a stressed syllable in the NORMAL condition was compared across conditions and referred to in all conditions as vowel in a stressed syllable.

Total durations of the analyzed phrases, i.e., the sum of the durations of the acoustic segments of each phrase, did not differ systematically in the SLOW SAF and DAF conditions. The acoustic segments were normalized as follows:

$$RD = \frac{D(Seg)}{\Sigma D(Seg)}$$

where, for a sentence or phrase under investigation, RD is the relative duration of a segment, D(Seg) is the duration of an acoustic segment in that sentence or phrase, and $\Sigma D(Seg)$ is the sum of durations of the acoustic segments in that sentence of phrase. Briefly, RD simply represents the relative duration in percent of a certain segment in an utterance for a specific condition.

The so normalized data revealed that compared to the NORMAL condition,

(1) the vowels were of longer relative duration in both the SLOW SAF and DAF conditions;

(2) the stops were of shorter relative duration in both the SLOW SAF and DAF conditions.

Compared to the SLOW SAF condition,

(1) vowels in stressed syllables, except for the open vowels /æ/ and /ɔ/, were of shorter relative duration in the DAF condition;

(2) fricatives were of longer relative duration in the DAF condition.

Graphs of the relative durations of syllables revealed patterns which were similar for the NORMAL and SLOW SAF conditions in that word- and sentence-stress correlated with longer relative duration, whereas the patterns for the DAF condition were unlike those of the NORMAL condition in that final syllables of the sentence were of longer relative duration regardless of stress and in that fricatives accounted for the longer relative duration of syllables regardless of stress.

In summary, position in the syllable, stress, degree of vowel openness and manner of articulation appear to be parameters which determine changes at a slower rate of speech. In addition, voluntarily slowed speech appears to differ from DAF-slowed speech. Since the results of the

pilot study are based on the performance of one speaker only and since no attempt was made to control rate under DAF, they were considered merely hypotheses to be tested.

6.2 Screening Study

The two different ranges of delay with which the two sets of subjects were presented had been chosen on the basis of previous findings (Mahaffey and Stromsta, 1965), that the delay producing maximum disturbance was 180 msec for males and 270 msec for females. The graphs of the duration-error ratio described in chapter 4, revealed that for each subject there occurred on the interval measured at least two high values, as calculated by the above-mentioned index of disturbance, separated from each other by a lower value. The delay producing the higher of the two high values was systematically chosen as the delay (producing maximum disturbance) for the main experiment. The results are summarized in table I. As can be seen, the delays producing the highest disturbance values tended to be shorter for males than for females. The average of the delays producing the highest value fell near 200 msec. The average of the delays producing the second high value was 297 for females and 206 for males.

Order of delay presentation may have influenced which delay produced maximum disturbance for male subjects 3, 4, and 5. For these three subjects the first reading under DAF produced the maximum disturbance. No such order

Subjects	Reading Time in seconds under SAF	Delay in msec producing highest peak disturbance	Delay in msec producing the second peak
Female subjects			
1	25	250	290
2 (a)*	24	200	290
2 (b)	24	200	310
3	25.5	230	310
4	23.5	200	310
5	22.5	200	270
Male subjects			
1	23.5	240	200
2	28	240	170
3	21.5	170	220
4	28	180	240
5	29	170	200

* female subject number 2 was recorded twice at different times with the delays presented in different orders to check if the measure was reliable.

Table I. Normal reading time under SAF and delays producing peaks of disturbance as defined by the duration-error ratio described in chapter 4.

effects were observed for the remaining subjects. Timmons (1971) observed order effects for females, whereas she observed no order effects for males. The results of the screening study suggest that perhaps order effects are not entirely confined to females.

Figure 6.1 shows the relative amount of disturbance for the three subjects whose recordings were analyzed in the main experiment. As can be seen, each of the three subjects was affected by DAF to a very different extent. Many hypotheses have been offered in the past to account for such individual differences. Beaumont and Foss (1957) found that subjects, who had high scores on tests of perseveration, were more affected by DAF. These authors related personality variables to degree of DAF disturbance. Zelniker (1971) and Salter (1973), among others, suggested that attentional factors determine the extent to which a subject is affected by DAF, e.g., if either another task demands attention simultaneously or if the DAF task (reading material) is more difficult (random words as opposed to prose), DAF will disrupt the speech less. Yates (1965) hypothesized that the factor which determines degree of speech disturbance under DAF is the extent to which a speaker relies on airconducted auditory feedback for normal control of speech, assuming that some speakers rely on the different feedback channels for control of speech to differing extents.

In summary, the results of this part of the study confirm the considerable individual differences which exist

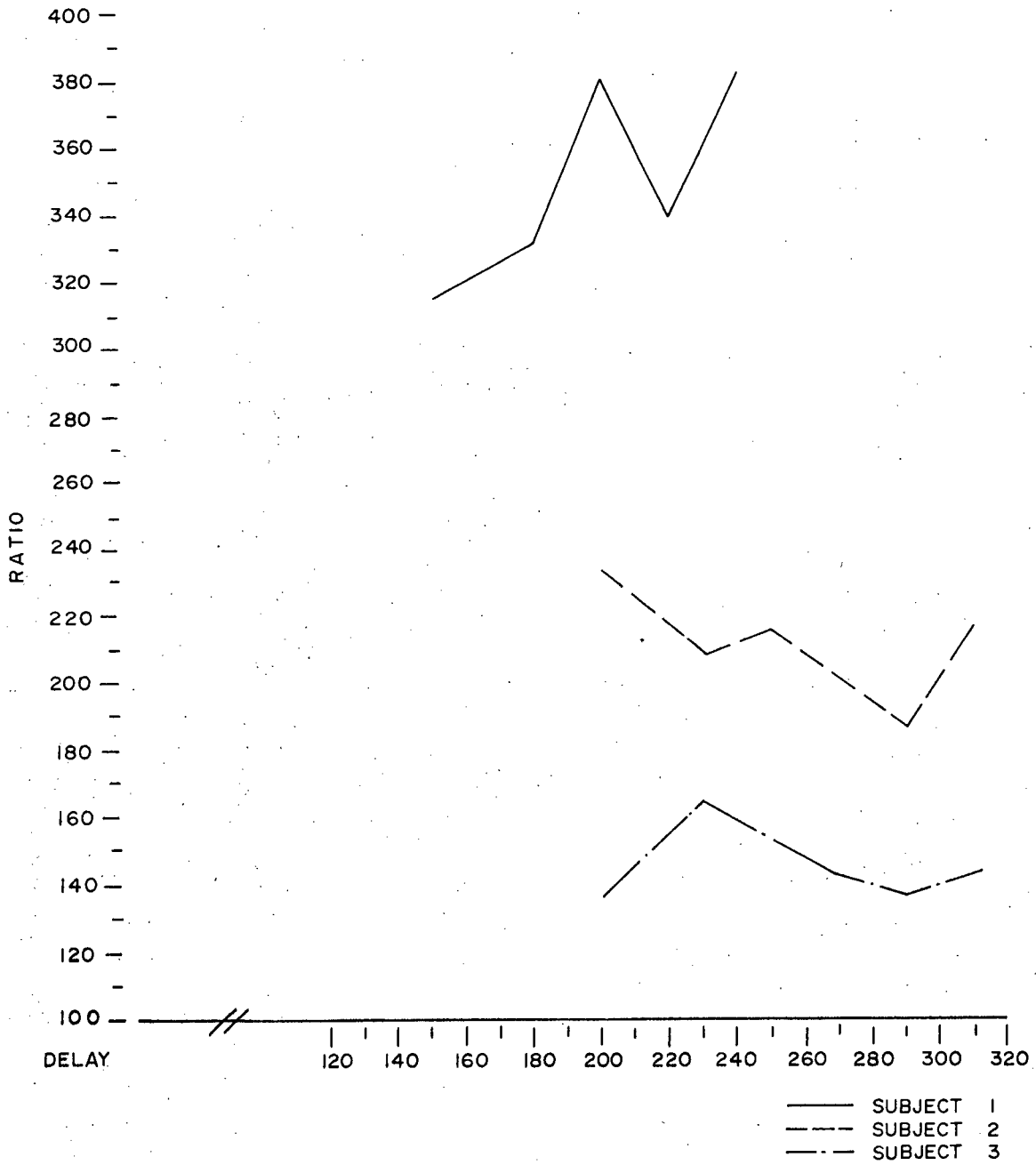


Figure 6.1. DAF-induced speech disturbance (error-duration ratio) for male subject 1, and female subjects 2 and 3.

in reaction to DAF, and, although the screening study was not designed to answer questions about the nature of DAF interference or about the nature of speech production units, the results suggest, as Chase et al. (1959) have suggested, that the critical time relationships are perhaps the most important factors to be examined if the observed changes in speech under DAF are to be explained.

6.3 Results of the Main Experiment

In accordance with the criteria listed in chapter 5, the mingograms of the main experiment passage in the four conditions were segmented for male subject 1 and the mingograms of the main experiment sentences in the four conditions were segmented for male subject 1 and female subjects 2 and 3. The duration of each acoustic segment was measured for every segmented utterance. 'Total duration' of a sentence or phrase was defined as the total duration of the utterance minus the pauses and minus the rejected segments (according to the rejection criteria outlined in chapter 5). Figure 6.2 shows the average 'total durations' of the four sentences in the four conditions for the three subjects. 'Total durations' varied from subject to subject and from condition to condition. In order to evaluate if the same speaking rate was maintained throughout a sentence, the absolute durations of each syllable divided by the number of phonemes in that syllable were graphed (see Appendix G for sample graphs). The graphs revealed that subject 1

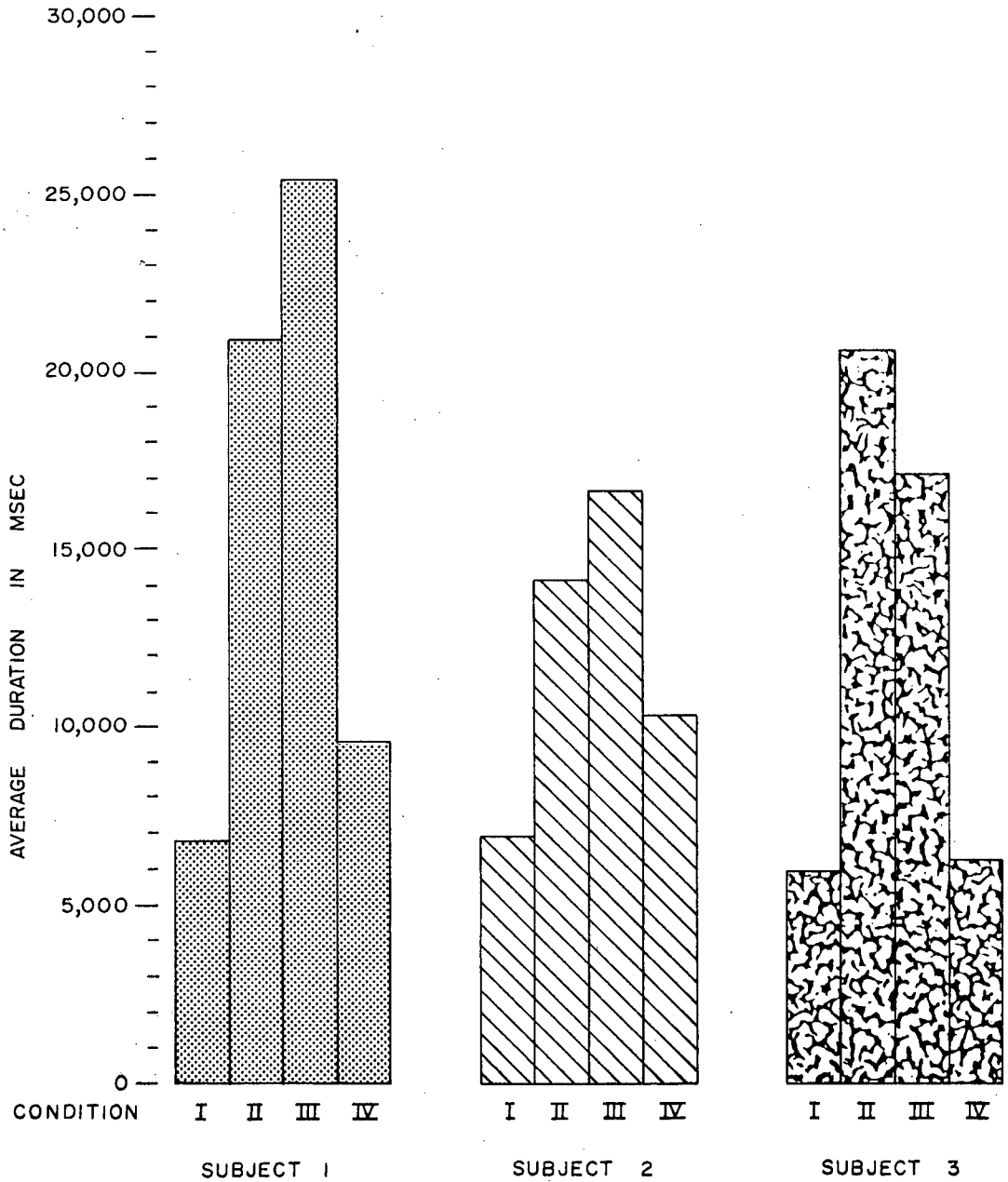


Figure 6.2. Average 'total durations' of the main experiment sentences in the four conditions.

tended to maintain the same speaking rate throughout a sentence, while subjects 2 and 3 tended to increase their rate towards the end of sentences in the SLOW SAF and SLOW DAF conditions. The 'total duration' of each phrase in the passage revealed that during the SLOW SAF condition subject 1 tended to slow down towards the end of the passage. In short, the speaking rates varied across subjects and conditions and at times within conditions.

Although subjects spoke at their maximum rate in the FAST DAF condition as opposed to speaking at their normal rate in the NORMAL condition, it will be noted that 'total duration' for all subjects under FAST DAF was greater than in the NORMAL condition, even though no repetitions or other additions to the expected sequence of articulation were included in the 'total duration' values as plotted in figure 6.2. If the durations of the repetitions are added to the 'total duration' for each subject, subject 1 exhibits by far the most slowing under DAF and subject 2 exhibits more slowing than subject 3. These facts are mentioned here for two reasons. Firstly, the data support the error-duration ratio as a reliable indicator of degree of speech disturbance produced by DAF in that the same results were obtained by the error-duration ratio. Secondly, the data seem to indicate that repetitions and prolongation are indeed manifestations of the <u>same</u> phenomenon, namely slowing: subjects who under FAST DAF do not make many repetitions do not prolong very much. This does not answer why a

repetition occurs at one time and drawling at another, it merely suggests the similarity of the effects in terms of subject reaction.

6.3[1] Relative Durations

The data were normalized as described in the pilot study. The relationships of 'stressed' long vowels as defined below under 6.3[3], vowels other than the 'stressed' long vowels ('other'vowels) , and consonants are shown in figure 6.3 for the main experiment sentences in the four conditions. As can be seen, in comparison with the NORMAL condition, vowels, primarily 'other'vowels tended to increase in relative duration and consonants tended to decrease. No consistent change in relative duration occurred for /á/ and /í/ from condition to condition. Compared to the other conditions, /ó/ tended to increase in relative duration in the FAST DAF condition. In comparison with the other conditions,/ú/ consistently increased in relative duration in the SLOW SAF condition.

A discussion of these results, which are essentially the same as the results of the data analysis described below, will follow when the latter results are reported and discussed.

6.3[2] Relative Change Measure

The data were also normalized as follows:

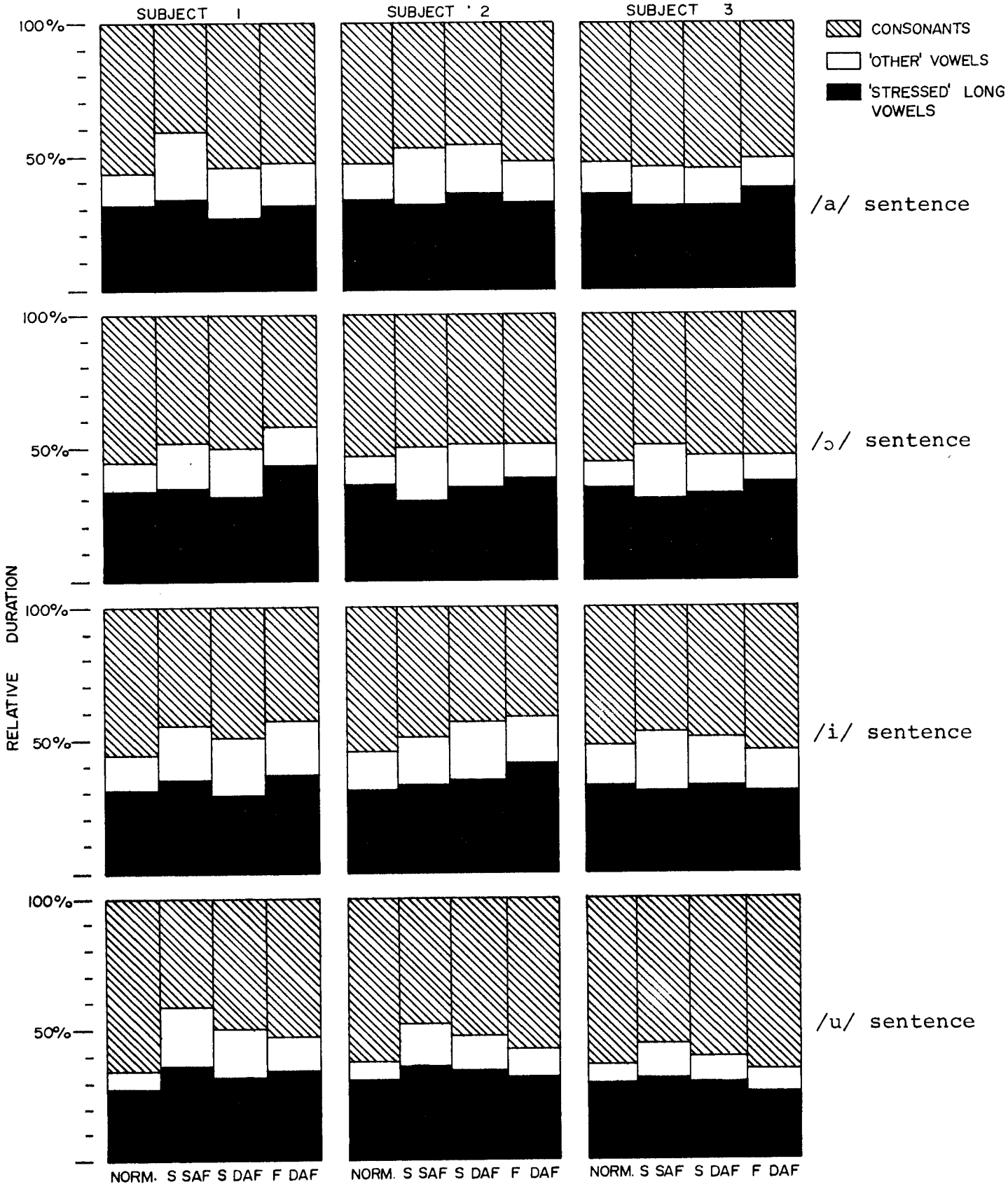


Figure 6.3. Relationship of 'stressed' long vowels, 'other' vowels, and consonants in the main experiment sentences in the four conditions.

$$RC(k) = \frac{SD(k)}{SD(I)} \times \frac{TD(I)}{TD(k)}$$

where, for a particular segment, RC(k) is the relative increase or decrease in duration of an acoustic segment in a particular condition k(II, III, or IV) with respect to the corresponding segment in the NORMAL condition, SD(k) is the duration of the acoustic segment under condition k, and TD(k) is the 'total duration' under condition k. This measure reflects the relative increase or decrease in duration of an acoustic segment compared to its duration in the NORMAL condition and normalized with respect to the two corresponding 'total durations'. An RC value of 1.0 corresponds to no change, an RC value greater than 1.0 corresponds to a relative increase and an RC value less than 1.0 corresponds to a relative decrease. 'Increase' or 'decrease' will be used hereafter to refer to the relative increase or decrease calculated by means of the above formula.

The RC values obtained from the analysis of the four main experiment sentences will be reported for different sets of phones separately.

Any levels of significance mentioned in this study will refer to levels of significance established by t-tests as follows:

* = significant beyond the 0.05% level

** = significant beyond the 0.02% level

*** = significant beyond the 0.01% level

**** = significant beyond the 0.005% level

6.3[3] Relative Change of Vowels

It became necessary in the course of analysis, on the basis of their apparently different functions, to consider vowels in monosyllabic function words such as prepositions, pronouns, and determiners, separately from the vowels in other words, primarily words bearing lexical information. Hereafter 'stressed' vowel will refer to vowels in stressed syllables in words other than the monosyllabic function words, and 'unstressed' vowel will refer to vowels in all other syllables. 'Unstressed' long and short and 'stressed' long and short vowels were analyzed separately. Figure 6.4 displays the relative change for the different groups of vowels for each subject. As can be seen, the vowels in the function words 'increased' more than any other group of vowels. Table II shows the between-condition statistical differences for vowels in function words and table III shows the within-condition statistical differences. The 'increases' for vowels in function words in the SLOW SAF condition were significantly different from the 'increases' in the FAST DAF condition and to a lesser extent from the 'increases' in the SLOW DAF condition. Furthermore, the 'increases' for the vowels /u/, /ɔ/, and /ə/ in function words were frequently significantly different from the 'increases' for /ú/, /ó/ and 'stressed' vowels respectively. Most of the significant differences occurred in the SLOW conditions. The relative changes for the vowels /ʌ/ and /æ/ in function words were rarely significantly

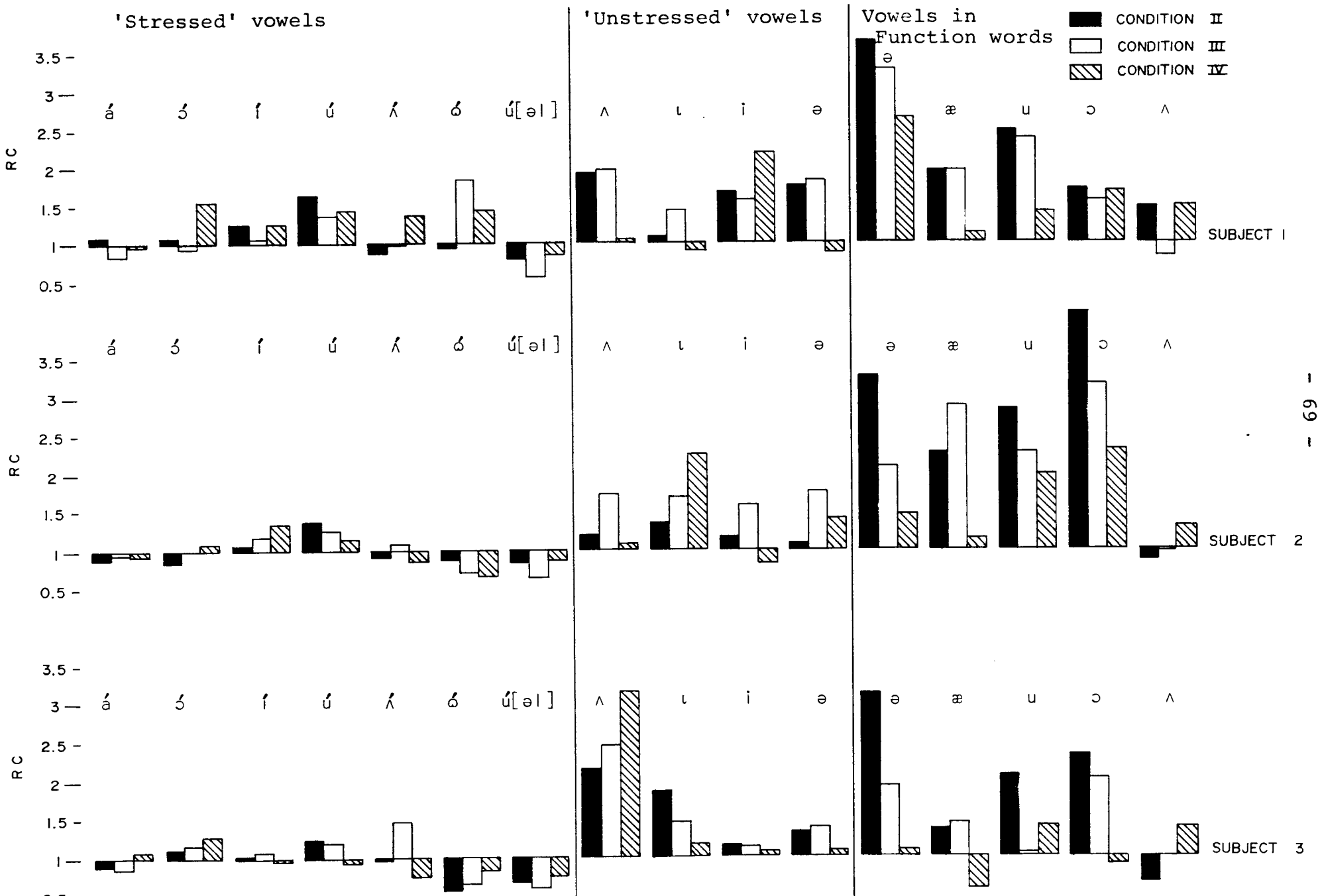


Figure 6.4. Relative change in duration for vowels in the SLOW SAF (II), the SLOW DAF (III), and the FAST DAF (IV) conditions with respect to the NORMAL condition.

		SLOW DAF	FAST DAF
	Subject 1	NS	**
SLOW SAF	2	*	****
	3	**	****

Table II. Between-condition differences in relative change of vowels in one-syllable function words.

	SLOW SAF	SLOW DAF	FAST DAF
/ú/ vs. /u/ in function words, /u/ sentence			
Subject 1	*	****	NS
2	****	****	**
3	***	NS	***
/ɔ́/ vs. /ɔ/ in function words, /ɔ/ sentence			
Subject 1	NS	***	NS
2	****	****	****
3	*	NS	NS
/á/ vs. /ʌ/ and /æ/ in function words, /a/ sentence			
Subject 1	****	NS	NS
2	NS	*	NS
3	NS	NS	NS
'stressed' vowels vs. /ə/ in function words, all sentences			
Subject 1	****	****	****
2	****	****	NS
3	****	****	NS

Table III. Within-condition differences in relative change of 'stressed' long vowels and vowels in function words.

different from the relative changes for the vowel /á/. The small number of occurrences of /ʌ/ and /æ/ (N=3) might account for the nonsignificant results for these vowels. However, an alternative explanation will be offered below, when these results are discussed. In order to assure that the difference between 'stressed' vowels and /ə/ in function words was not simply due to the differences in vowel quality, /ə/ in words other than function words, e.g., in /sitəd/, were compared with 'stressed' vowels. The statistical results, presented in table IV, indicate that there is rarely a significant difference in relative change between /ə/ in words other than function words and 'stressed' vowels.

Allen (1968) in a discussion of rhythmic behaviour in motor tasks proposed that the minimum interval between acts for rhythmic behaviour in speech (beats) is probably 300 msec and the maximum interval 1000 msec. It was found in the present study that when the vowels in function words are examined, the speech power trace on the mingograms is as high for these vowels as for the vowels which have been called 'stressed' vowels in this study. When the data are examined in view of Allen's minimum and maximum limits, it becomes evident, that the subjects, at the slower rates of speech and particularly at the very slow rate in the SLOW SAF and SLOW DAF conditions, lengthened vowels and increased the physical effort with which they were produced in comparison with the NORMAL condition such, that each vowel became a beat in the rhythmic succession of speech

	SLOW SAF	SLOW DAF	FAST DAF
'Stressed' vowels vs. /ə/ not in function words			
Subject 1	NS	NS	NS
2	NS	***	NS
3	**	*	*

Table IV. Within-condition differences in relative change of 'stressed' vowels and /ə/ not in function words.

units. Further examination revealed that the intervals between the normally stressed vowels would have exceeded 1000 msec at the slower rates of speech, thus necessitating the additional beats. In other terms, the slower the rate of speech, the more likely will a normally unstressed vowel be prolonged and form a beat in the rhythmic succession of speech units. This may explain why the differences between the SLOW SAF and SLOW DAF conditions for vowels in function words were not as significant as the differences between the SLOW SAF and FAST DAF condition and why in the FAST DAF condition there should be fewer significant differences between 'stressed' vowels and vowels in function words. The exceptions, i.e., the vowels in function words which did not increase significantly in duration in comparison with other vowels, were now examined in view of the above hypothesis: it was found that these exceptions occurred between two beats (two stressed vowels) which were not separated from each other by more than 1000 msec. In addition, these exceptional vowels, e.g., /ʌ/ in /ʌp/ and /æ/ in /æt/, occurred in an environment that in the NORMAL condition would call for a shorter vowel. A discussion of this latter point follows when the results of the analysis of 'stressed' vowels are discussed.

Analysis of the 'stressed' vowels (see figure 6.4) revealed that in general /ú/ was consistently most 'increased' in the SLOW SAF condition, and more 'increased' in the SLOW DAF condition than in the FAST DAF condition.

The reverse was true for /ɔ́/, where the greatest increase occurred in the FAST DAF condition. /ú/ was more 'increased' than any other 'stressed' vowel in any other condition with the exception of /ɔ́/ in the FAST DAF condition. /ú/ in /úəl/ was not included in the analysis of /ú/ because of the consistently different behaviour of /ú/ in this environment. Table V shows the within-condition differences in relative change for the four groups of 'stressed' vowels. As can be seen, the 'increases' for /ú/ were significantly different from 'increases' for the low vowels (primarily /á/) in the SLOW SAF and SLOW DAF conditions for subjects 1 and 2. For subject 3 this difference rarely reached significance, although the direction of the difference was the same, namely /ú/ was more 'increased' than the low vowels. The only significant between-condition differences for 'stressed' vowels (not shown in table form) were as follows: with respect to the SLOW SAF condition, the 'increases' for /ɔ́/ and /í/ were each significantly different for one subject in the FAST DAF condition and /á/ was significantly different for one subject in the SLOW DAF condition.

Analysis of the vowels in the passage substantiated the above findings in that vowels in function words 'increased' more than any other type of vowel and most at the slowest rate, i.e., in the SLOW SAF condition. The average RC values for vowels in the passage amounted to 1.51 in the SLOW SAF condition, 1.38 in the SLOW DAF

		/á/	/ɔ́/	/í/	/ú/
SLOW SAF					
/á/	Subject 1	-	NS	NS	****
	2	-	NS	NS	****
	3	-	NS	NS	NS
/ɔ́/	Subject 1	-	-	NS	**
	2	-	-	*	****
	3	-	-	NS	NS
/í/	Subject 1	-	-	-	NS
	2	-	-	-	*
	3	-	-	-	NS
SLOW DAF					
/á/	Subject 1	-	NS	NS	****
	2	-	NS	NS	***
	3	-	NS	NS	***
/ɔ́/	Subject 1	-	-	NS	**
	2	-	-	NS	*
	3	-	-	NS	NS
/í/	Subject 1	-	-	-	**
	2	-	-	-	NS
	3	-	-	-	NS
FAST DAF					
/á/	Subject 1	-	NS	NS	****
	2	-	NS	***	NS
	3	-	NS	NS	NS
/ɔ́/	Subject 1	-	-	NS	NS
	2	-	-	*	NS
	3	-	-	NS	NS
/í/	Subject 1	-	-	-	NS
	2	-	-	-	NS
	3	-	-	-	NS

Table V. Within-condition differences in relative change of 'stressed' vowels.

condition, and 1.20 in the FAST DAF condition.

Why was /ú/ more 'increased' than the other vowels and particularly more than /á/ in the SLOW conditions? Two different answers will be examined in detail: (a) the vowel /u/ itself differs from the other vowels such that it explains the greater 'increase', (b) the environment in which /u/ occurred, differed from the environment in which the other vowels occurred such, that it will explain the greater 'increase'.

(a) The vowel /u/ differs, as does /i/, from /a/ and /ɔ/ in that it has a lower intrinsic amplitude (Lehiste and Peterson, 1959) and /u/ differs from all other vowels in that it is close and back. Why /u/ should behave differently cannot be explained in terms of these facts.

(b) House and Fairbanks (1953) found that voicing, manner and place of articulation, in that order, exert an influence on vowel duration, e.g., in voiced environments vowels are generally longer; when preceded and followed by fricatives, vowels are longer, whereas when preceded and followed by stops, vowels are generally shorter; when preceded and followed by consonants with a bilabial place of articulation, vowels are generally shorter. According to Peterson and Lehiste (1960), the influence of the preceding consonant on vowel duration is minimal. Chen (1970) confirmed that the increase in vowel duration before a voiced consonant is true for other languages as well. Delattre (1962) suggested that the changes in vowel duration are

for the most part physiologically conditioned, i.e., "anticipation of greater effort" (clusters, closures, no voicing) "would make one 'shorten' the vowel" (ibid., p. 1142). Chen (1970) proposed that observed faster transitions from an open vocal tract (vowel) to a more close position for a voiceless consonant explain most satisfactorily the shorter duration of a vowel before a voiceless stop.

When the /u/ sentence is now compared with the other sentences, it is evident that /ú/ appeared more frequently in an environment which results in a shorter duration of the vowel according to the above findings. The vowel occurring least frequently in such an environment was /á/. Does this imply that, when the rate of speech is decreased, vowels in an environment, which results in a longer vowel duration, are prolonged relatively more than vowels in an environment which shortens the duration of the vowel? In other terms, do longer segments at slower rates of speech tend to be relatively more prolonged than shorter segments?

The other fact of interest emerging from the analysis of 'stressed' vowels was the consistently greater 'increase' for /ɔ́/ in the FAST DAF condition. In view of the earlier reported findings by Sussman and Smith (1971) that jaw activity for the vowels /i/, /ɛ/ and /æ/ was longest at 100-msec delay, it is perhaps not too surprising that in the present study, where longer than 100-msec delays were used, no consistent 'increases' under DAF for the

vowels /í/, /ú/ and /á/ were found. How, then, can the consistent 'increases' for /ó/ be explained? Is it possible, that for /ó/ a longer than 100-msec delay is needed to produce longest jaw activity? Any hypothesis which would attempt to explain selective vowel duration under DAF would have to account for the distance the articulators have to move from the rest position as well as for critical time delays. In the present study, vowel context, which turned out to be a possible source of variation of 'increases' for vowels, was not controlled well enough, so that no such hypothesis can be offered.

6.3[4] Relative Change of Consonants

In general, consonants tended to be 'decreased' in the SLOW SAF condition, whereas vowels showed an 'increase'. The effect was statistically significant beyond the 0.01% level with one exception for the three subjects. Consonants tended to be 'decreased' also in the SLOW DAF and FAST DAF conditions but not as consistently or as significantly as in the SLOW SAF condition. This was also true for the analyzed main experiment passage.

It has been suggested when the greater increases for /ú/ in the SLOW conditions were discussed, that an environment which results in a longer segment duration, also appears to result in greater 'increases' at slower rates for that segment. The question was asked whether longer segments at slower rates of speech tend to be

relatively more prolonged than shorter segments. If this is in fact the case then it would seem reasonable that consonants, in particular noncontinuants, should be 'decreased' or less 'increased' in comparison with vowels.

Under DAF a number of other factors must be considered. A discussion follows after the results of the consonant analysis are reported. In the statements below, 'initial' and 'final' will refer not necessarily just to the first or last consonant in the syllable but will mean pre- or postvocalic respectively. Figure 6.5 shows the average RC for 'initial' closure and 'final' stops, and for 'initial' /m/ and 'final' nasals. As can be seen, 'initial' closure and 'initial' /m/ either 'increased' more under the DAF conditions or 'decreased' less than in the SLOW SAF condition. 'Final' nasals consistently 'increased' in the SLOW SAF condition and either 'increased' less or 'decreased' in the DAF conditions. As can be seen, no consistent pattern emerged for the 'final' voiceless stops. It will be remembered that 'initial' closure refers in this study to the closure part of an 'initial' stop only, whereas 'final' stop includes the release of the closure and aspiration. Analysis of the passage revealed that 'initial' /b/ was also increased under the DAF conditions, the increase in the DAF conditions being highly significantly different from the SLOW SAF condition. In the sentences, where no such observation was made, /b/ occurred only after a consonant prevocalically, whereas in

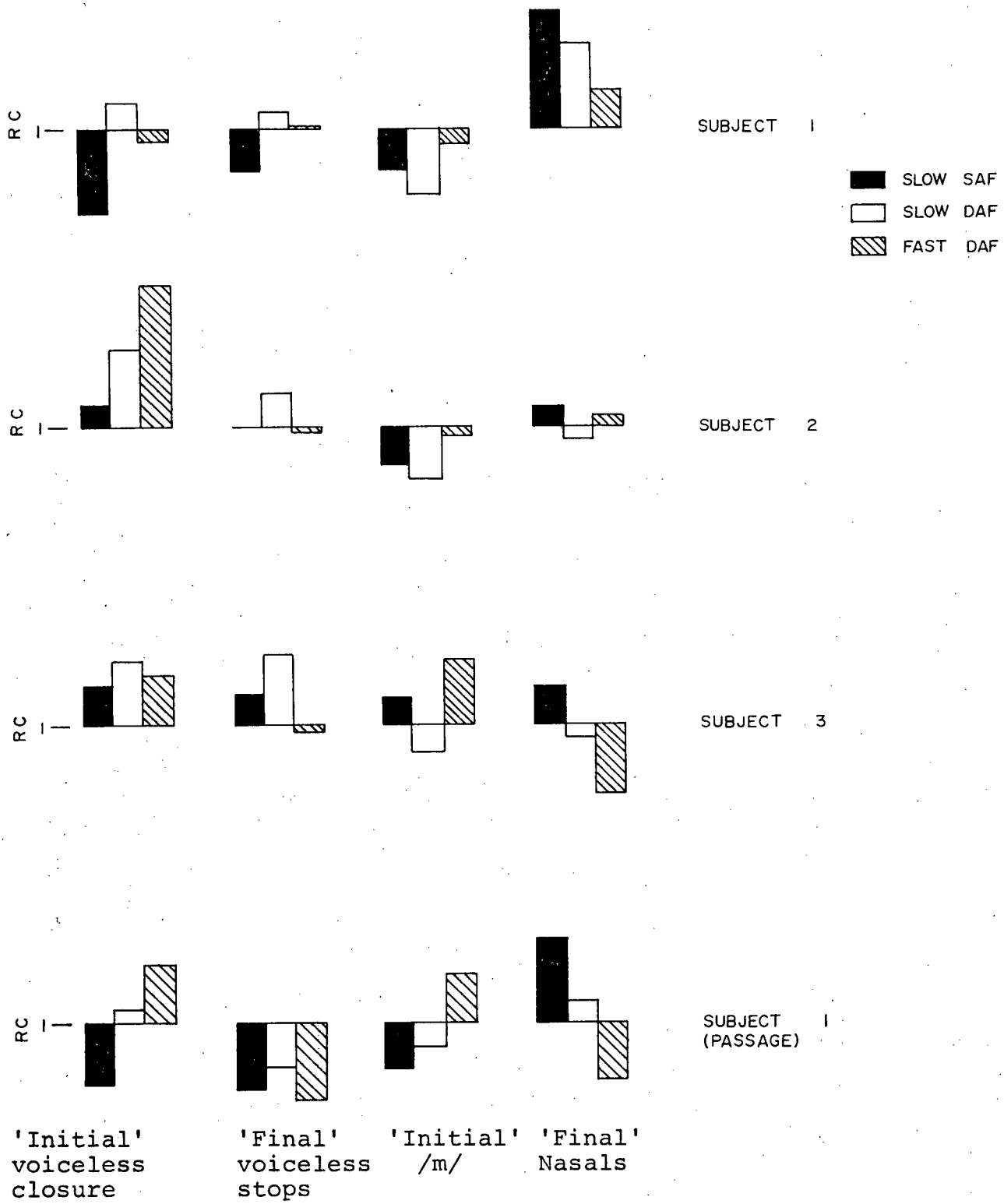


Figure 6.5. Relative change in 'initial' voiceless closure, 'final' voiceless stops, 'initial' /m/ and 'final' nasals.

the passage /b/ occurred only word-initially. Figure 6.6 shows the relative change for aspiration associated with the release of 'initial' voiceless stops. As can be seen, aspiration 'decreased' for all subjects in all conditions. However, when between-condition differences in relative change for aspiration are examined, it can be seen in Figure 6.6 that subjects 1 and 2 showed an even further 'decrease' in aspiration in the DAF conditions, whereas subject 3 showed a lesser 'decrease'. Table VI shows the statistical significance of these observations. Subject 3 differed further from subjects 1 and 2 in that fricatives were either more 'increased' or less 'decreased' in the DAF conditions than in the SLOW SAF condition; (significant at the .05% level in the SLOW DAF condition and at the .005% level in the FAST DAF condition). The differences, although in the same direction for subjects 1 and 2, did not reach significance for these subjects.

Haggard (1971), as quoted in MacNeilage (1972), found that of the postvocalic consonants, the absolutely final consonants appear to have the longest durations. In the present study, consonants in this position were frequently more 'increased' than consonants in other positions in the SLOW SAF condition. 'Final' nasals illustrate this point, in that for all subjects the highest RC values for 'final' nasals were obtained consistently in absolutely final position. This appears to further confirm the impression that at a slower rate, the normally longer

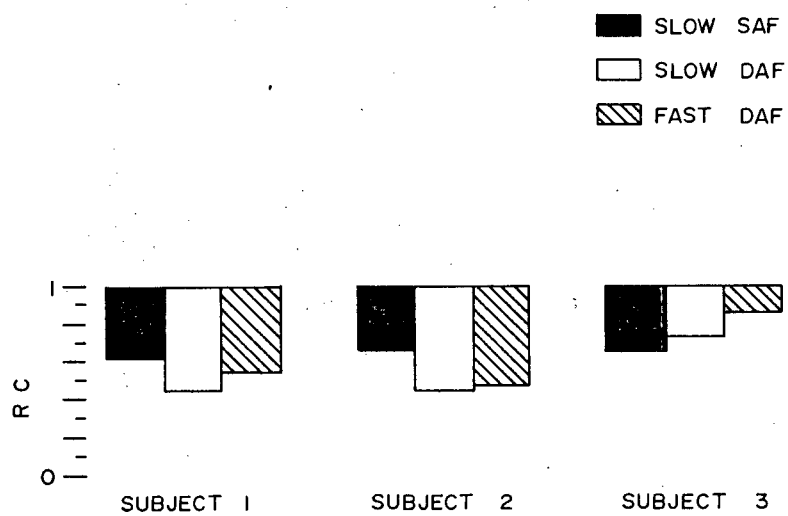


Figure 6.6. Relative change in aspiration in the SLOW SAF, SLOW DAF and FAST DAF conditions with respect to the NORMAL CONDITION.

		SLOW DAF	FAST DAF
SLOW SAF			
	Subject 1	*	NS
	2	****	*
	3	NS	*

Table VI. Between-condition differences in relative change of aspiration after 'initial' voiceless closure.

segments are 'increased' more than normally shorter segments, or 'decreased' less. If this impression is true, then it would be expected that under the SLOW SAF condition closure would exhibit the lowest RC value. This was not always the case, particularly not for voiced closure. Furthermore, of the different closures one might expect the bilabial closure to show a higher RC value under the SLOW SAF condition than dental or velar closures because, as MacNeilage (1972) has pointed out, bilabial closure can be maintained longer than dental or velar closure since the tongue in the latter closures is needed in the next articulation. However, the expected systematic differences in relative change for the different closures were not found. It appears then, that, while it is true that in the SLOW SAF condition vowels are more'increased' than consonants and vowels in an environment, in which the vowel is longer, more 'increased' than vowels in an environment, in which the vowel is shorter, and absolutely final consonants more 'increased' than consonants in other post-vocalic positions, it is not true that <u>all</u> normally longer segments are more 'increased' under the SLOW SAF condition. Instead, position in the syllable, voicing, and place and manner of articulation, beyond a simple stop / continuant dimension for consonants, all appear to play differential roles in determining 'increases' for consonants in the SLOW SAF condition. In order to specify the role of each of the parameters, environment would have to be strictly

controlled.

The findings of the consonant analysis are best discussed, however, in terms of reaction to DAF. To summarize, 'initial' voiceless closure, 'initial' voiced closure (not in a cluster), 'initial' /m/ and fricatives showed a greater 'increase' or a lesser 'decrease' in the DAF conditions than in the SLOW SAF condition. Further analysis revealed that bilabial closure contributed the most to the observed differences involving closure, in that bilabial closure exhibited most consistently a higher RC value under DAF than under SLOW SAF. It must be pointed out here, that the observed differences in relative change for 'initial' closure cannot be explained in terms of consistently lower RC values for closures under the SLOW SAF condition, but instead the differences must be attributed to the DAF effect. It will be remembered that Huggins (1968) found that initial voiceless closure grew linearly with delay. 'Initial' position appears to be a crucial parameter in determining 'increases' under DAF. MacNeilage (1972) in a discussion of closed loop control system in speech concludes that one cannot assume that auditory feedback plays an important role in the control of running speech, but somatic sensory control is considered necessary for the initiation of speech.

> "The command necessary for an articulator to reach a fixed speech-initial position must be conditional upon the pre-speech position of that articulator which no doubt varies from occasion to occasion." (ibid.,pp.44-45).

It is entirely possible, that in the normal case this is the only type of feedback needed for the initiation of speech. If the system under DAF, however, receives the normal somatic sensory feedback with respect to articulator position and at the same time receives air-conducted auditory feedback which informs the system that it (the air-conducted auditory feedback) is at variance with the somatic sensory information, then hesitation might be expected. If we furthermore assume that once speech is initiated an open loop system takes over, than it is possible that under DAF the beginning of a speech unit should be prolonged relatively more than would be expected for a slower rate of speech. This does not explain the greater 'increases' for vowels. However, since vowels were more 'increased' also at the slower rate of speech under SAF, this phenomenon cannot be attributed to DAF but must be considered typical of a slower rate of speech in general.

6.3[5] Repetitions

Figure 6.7 shows the number of repetitions of different durations for each subject in the FAST DAF condition and the cumulative % repetitions. The graphs revealed that length of the repeated segments, as defined in chapter 5, correlated with the length of delay under which they were produced. As can be seen, subject 1 repeated more frequently than subject 2 and subject 2 more frequently than subject 3. It might be mentioned here that the latter

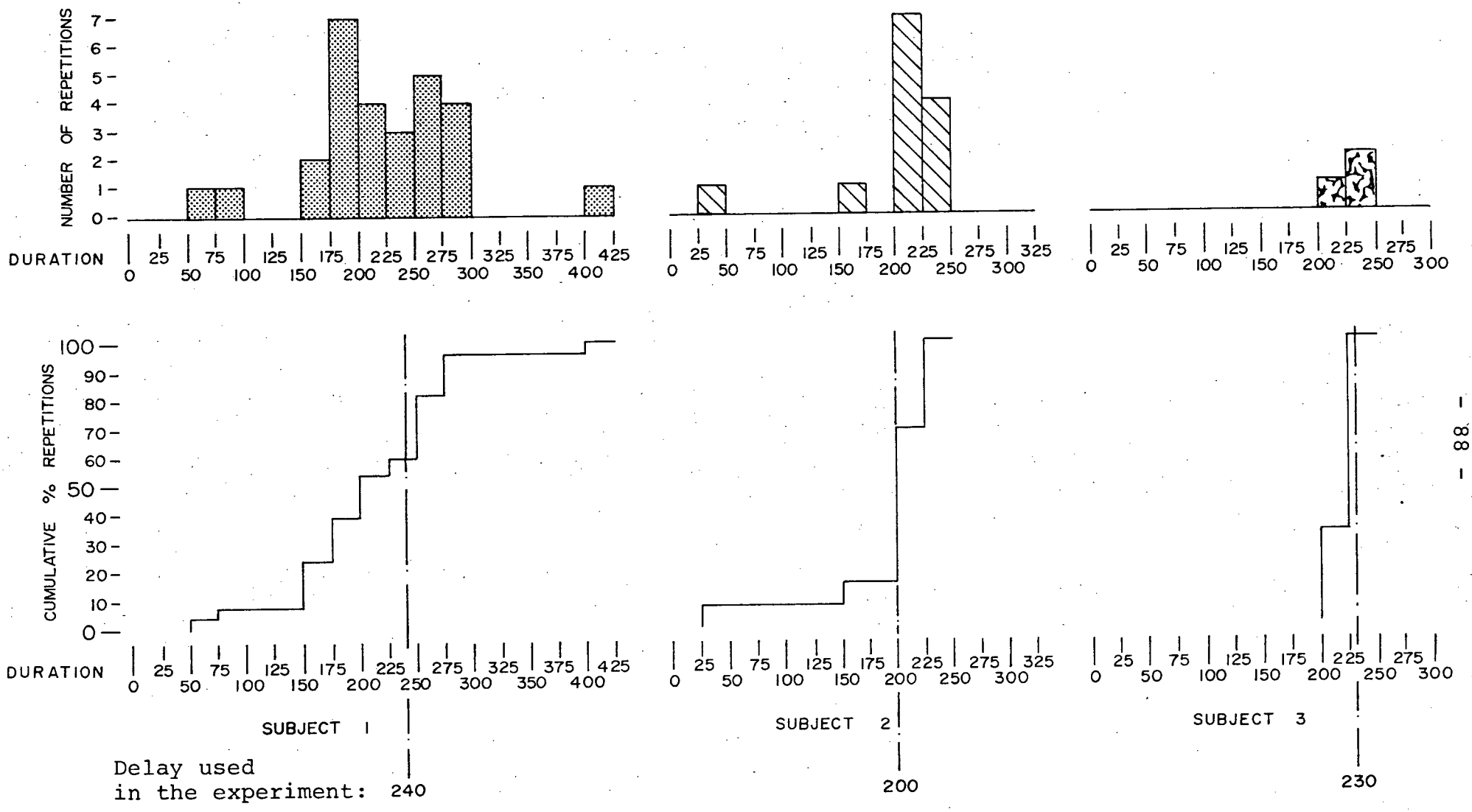


Figure 6.7. Number of repetitions and cumulative % repetitions in the FAST DAF condition.

finding was also true for other additions to the expected sequence of articulation. Subject 1 exhibited by far the most additions unrelated to the text and rather bizarre in nature, e.g., /kam faðɚ/ became /k^{h}ambagl̩ feβafaðɚ/. In the SLOW DAF condition only a few isolated repetitions occurred which is in agreement with MacKay's (1968) finding that under DAF at a slower rate repetitions decrease in number. In view of the rarity with which multiple repetitions have been observed for English speakers under DAF by past experimenters, it might be noted here, that although subjects 2 and 3 primarily exhibited double repetitions, subject 1 frequently produced multiple repetitions (up to five repetitions in one instance). This does suggest that multiple repetitions are not peculiar to the Russian speakers in the Kozhevnikov and Chistovich (1965) study. Furthermore, these repetitions were always repetitions of open syllables, which is in agreement with the findings of the above authors. Finally, it was observed that when a vowel was repeated at the beginning of a syllable, e.g., /iʔist/, a glottal stop was inserted, and when a vowel was repeated after a consonant, e.g., /kaham/, an /h/ was inserted. Laryngealization frequently preceded a vowel in syllable-initial position in the DAF conditions.

Stromsta (1959) found that side-tone, which was distorted by altering its frequency components or its arrival phases at the two ears, produced a blockage in phonation while phonating a vowel in falsetto. Van Riper

(1971), who, like many others (for a review see Timmons and Boudreau, 1972), consider natural and artificial stutter to be similar behaviours, hypothesized that asynchronies in timing of antagonistic muscle groups at the laryngeal level are responsible for the observed stutter. In support of such a theory, Freeman and Ushijima (1974) on the basis of their experiments, state:

> "fluent utterance is characterized by precise balance and timing of laryngeal abductor and adductor forces. In fluent utterance abductor and adductor forces act reciprocally, whereas, in stuttered utterance, this reciprocity is disrupted. In many stuttered utterances the lateral cricoarytenoid-generally presumed to be an adductor, with the specific function of medial compression-achieved abnormally high levels of activity." (ibid., p. S80).

The laryngealization and insertion of glottal stops observed under DAF in this study and also the 'decrease' in aspiration under DAF after 'initial' voiceless closure would tend to confirm that the effects of DAF on speech are not unlike those of natural or artificial stutter. It is of particular interest in this respect that those subjects most affected by DAF showed a greater 'decrease' of aspiration under DAF than in the SLOW SAF condition, whereas the subject least affected showed a lesser 'decrease'. No explanation can be offered at this time for the differential insertion of glottal stops and /h/. Nor can the bizarre insertions into the expected sequence of articulation be reasonably explained. The correlation of the durations of repeated segments with the delay under which they were produced, is in

agreement with the findings reported by Huggins (1968).

In summary, the investigation of insertions into the expected sequence of articulation suggests that (a) mistiming at the laryngeal level accounts for some of the speech disturbances observed under DAF, (b) time relationships are a crucial aspect of behaviour under DAF in that the durations of repetitions are related to the delay under which they were produced and (c) multiple repetitions constitutes a behaviour which must in the last analysis be explained in terms of individual differences.

> "A major challenge that is emerging for researchers in speech physiology, when they use several subjects in their experiments, is to develop speech production theories which take these individual differences into account rather than ignoring them." (MacNeilage, 1972,pp. 22-23).

6.4 Summary and Conclusions

This study has examined and compared speech at a normal rate and at a slow rate under synchronous feedback and speech under DAF at both a slow and maximally fast rate.

It was found that considerable individual differences exist in terms of the degree of speech disturbance produced in the subjects under DAF. A reciprocal relationship between prolongation and repetition, i.e., the more overall prolongation, the fewer the repetitions, existed only when rate was varied under DAF. At a maximum rate, subjects who prolonged more also repeated frequently. In other terms, it was not true that the subject who repeated

most frequently, showed the least amount of overall prolongation. Repetitions at the maximum rate under DAF appeared to be related to the delay under which they were produced.

While the hypothesis that under SLOW SAF vowels are prolonged more than consonants was substantiated, it could not be confirmed that under SLOW SAF within the class of vowels, close vowels are prolonged more than open vowels. The hypothesis that under SLOW SAF continuant consonants are prolonged more than obstruents was also not substantiated.

The hypothesis that under DAF vowels and fricatives are prolonged the most and that position within the syllable affects the duration increases of each sound class selectively was confirmed.

The hypothesis that speech under SLOW SAF would differ from speech under DAF was substantiated in that under DAF open back vowels, fricatives, initial voiceless closure and all initial bilabial closure was relatively more prolonged then under SLOW SAF whereas under SLOW SAF consonants in final position were relatively more prolonged than under DAF.

It had also been hypothesized that stress would determine relative amount of prolongation under SLOW SAF. It was found that under SLOW SAF <u>and</u> under SLOW DAF and to a lesser extent under FAST DAF each syllable formed a beat in the rhythmic succession of speech units, i.e., there

are no unstressed segments at very slow rates of speech. The beat manifests itself in relative increases in duration of the vowel in the syllable and increased speech power in comparison with speech at a normal rate. Thus, normally unstressed segments, e.g., function words, at a very slow rate increase relatively more in duration than normally stressed segments.

In view of the discussion (Chapter 2), it can be reasonably assumed that slowing at a voluntarily slowed rate under SAF is not due to slowing at a reflex level, i.e., a peripheral level, but rather that such a voluntary slowing must take place at a central level, such as the motor control programme level proposed by Allen (1973). When the rate is slowed under DAF by complying with the DAF effect, slowing can be said to be voluntary as well. Thus, similarities in the resulting behaviour might be expected. At a maximally fast rate under DAF no voluntary slowing takes place. Any slowing must be attributed to the DAF effect. One might assume that the differences in speech behaviour under SLOW SAF and under DAF are attributable to differences in the level at which slowing takes place. If one assumes that voluntary rate changes are determined at a central level, than the observed differences between voluntarily slowed speech and DAF-slowed speech might be attributable to interference at the reflex level, while observed similarities might be attributable either to interference at a central level, or also to reflex level

interference, if reflex level interference is similar in these respects (from the point of view of results) to central interference. In any event, one cannot deny that the reflex level is involved in DAF-induced changes in the speech behaviour, especially in view of the involvement of the larynx as evidenced by frequent laryngealization under DAF. The prevocalic rather than postvocalic slowing observed under DAF and not under SLOW SAF seems to indicate that the syllable is a monitored unit in speech production under DAF; once speech is initiated, it is able to procede towards its conclusion, the next interference being noticeable again at the initiation of the next speech unit.

Although this study attempted to specify the relative changes in segment durations under DAF and attempted to outline those changes which are observed under DAF but not under SAF, it could not answer how these changes were caused beyond a confirmation of past experimenters' (Kozhevnikov and Chistovich, 1965) conclusions, namely that the reflex level is involved in the DAF-induced speech disturbances.

Bibliography

ALLEN, G.D. (1968). The Place of Rhythm in a Theory of Language. U.C.L.A. Working Papers in Phonetics No. 10.

ALLEN, G.D. (1973). "Segmental Timing Control in Speech Production," J. Phonetics 1, 219-237.

BEAUMONT, J.T., and FOSS, B.M. (1957). "Individual Differences in Reacting to Delayed Auditory Feed-Back," Brit. J. Psychol. 48, 85-89.

BLACK, J.W. (1951). "The Effect of Delayed Side-Tone upon Vocal Rate and Intensity," J. Speech Hearing Dis. 16, 56-60.

CHASE, R.A. (1958). "Effect of Delayed Auditory Feedback on the Repetition of Speech Sounds," J. Speech Hearing Dis. 23, 583-590.

CHASE, R.A., et al. (1959). "Comparison of the Effects of Delayed Auditory Feedback on Speech and Key Tapping," Science 129, 903-904.

CHASE, R.A., and GUILFOYLE, G. (1962). "Effect of Simultaneous Delayed and Undelayed Auditory Feedback on Speech," J. Speech Hearing Res. 5, 144-151.

CHEN, M. (1970). "Vowel Length Variation as a Function of the Voicing of the Consonant Environment," Phonetica 22, 129-159.

DELATTRE, P. (1962). "Some Factors of Vowel Duration and their Cross-Linguistic Validity," J. Acoust. Soc. Amer. 34, 1141-1143.

FAIRBANKS, G. (1954). "Systematic Research in Experimental Phonetics: 1. A Theory of the Speech Mechanism as a Servosystem," J. Speech Hearing Dis. 19, 133-139.

FAIRBANKS, G. (1955). "Selective Vocal Effects of Delayed Auditory Feedback," J. Speech Hearing Dis. 20, 333-346.

FAIRBANKS, G., and GUTTMAN, N. (1958). "Effects of Delayed Auditory Feedback upon Articulation," J. Speech Hearing Res. 1, 12-22.

FREEMAN, F.J., and USHIJIMA, T. (1974). "Laryngeal activity accompanying the moment of stuttering: a preliminary report of EMG investigation," J. Acoust. Soc. Amer. Supplement 55, S80.

FLETCHER, J.F., and YATES, A.J. (1971). "The Repetition of Single Speech Sounds under Delayed Auditory Feedback," Br. J. Psychol. 62, 73-79.

HOUSE, A.S., and FAIRBANKS, G. (1953). "The influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels," J. Acoust. Soc. Amer. 25, 105-113.

HUGGINS, A.W.F. (1968). "Delayed Auditory Feedback and the Temporal Properties of the Speech Material," Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung 21, 53-60.

KOZHEVNIKOV, V.A., and CHISTOVICH, L.A. (1965). Speech, Articulation, and Perception (translated from Russian), Joint Publication Research Service, U.S. Dept. Commerce No. 30 (Washington).

LADEFOGED, P. (1967). Three Areas of Experimental Phonetics (Oxford University Press, London).

LEE, B.S. (1950a). "Some Effects of Side-Tone Delay," J. Acoust. Soc. Amer. 22, 659-640.

LEE, B.S. (1950b). "Effects of Delayed Speech Feedback," J. Acoust. Soc. Amer. 22, 824-826.

LEE, B.S. (1951). "Artificial Stutter," J. Speech Hearing Dis. 16, 53-55.

LEHISTE, I., and PETERSON, G. (1959). "Vowel Amplitude and Phonemic Stress in American English," J. Acoust. Soc. Amer. 31, 428-435.

LENNEBERG, E.H. (1967). Biological Foundations of Language (John Wiley & Sons, New York).

MacNEILAGE, P.F. (1972). "Speech Physiology," in Speech and Cortical Functioning J. Gilbert, Ed. (Academic Press, New York).

MacKAY, D.G. (1968). "Metamorphosis of a Critical Interval: Age-Linked Changes in the Delay of Auditory Feedback that produces Maximal Disruption in Speech," J. Acoust. Soc. Amer. 43, 811-821.

MAHAFFEY, R.B., and STROMSTA, C.P. (1965). "The Effects of Auditory Feedback as a Function of Frequency, Intensity, Time, and Sex," De Therapia Vocis et Loquelae II, 233-235.

PETERSON, G., and LEHISTE, I. (1960). "Duration of Syllable Nuclei in English," J. Acoust. Soc. Amer. 32, 693-703.

PETERSON, G.E., and McKINNEY, N.P. (1961). "The Measurement of Speech Power," Phonetica 7, 65.

RAWNSLEY, A.I., and HARRIS, J.D. (1954). "Comparative Analysis of Normal Speech and Speech with Delayed Side-Tone by Means of Sound Spectrograms," Med. Res. Lab. Rep. No. 248, Bureau of Medicine and Surgery, Navy Dept. Proj. NM 003 041.56.03, 13, No. 9.

ROBINSON, G.M. (1972). "The Delayed Auditory Feedback Effect is a Function of Speech Rate," J. Exp. Psychol. 95, 1-5.

SALTER, D. (1973). "Capacity and Delayed Speech," Br. J. Psychol. 64, 193-198.

SMITH, K.U. (1962). Delayed Sensory Feedback and Behaviour (W.B. Saunders Co., Philadelphia & London).

STROMSTA, C.P. (1959.) "Experimental Blockage of Phonation by Distorted Sideton," J. Speech Hearing Res. 2, 286-301

SUSSMAN, H.M., and SMITH, K.U. (1970). "Jaw Movements under Delayed Auditory Feedback," J. Acoust. Soc. Amer. 50, 685-691.

TIMMONS, B.A. (1971). "Sex as a Factor Influencing Sensitivity to Delayed Auditory Feedback," Perceptual and Motor Skills 32, 824-826.

TIMMONS, B.A., and BOUDREAU, J.P. (1972). "Auditory Feedback as a Major Factor in Stuttering," J. Speech Hearing Dis. 37, 476-484.

VAN RIPER, C. (1971). The Nature of Stuttering (Prentice-Hall, Inc., Englewood Cliffs, N.J.).

YATES, A.J. (1965). "Delayed Auditory Feedback and Shadowing," Quart. J. Exp. Psychol. 17, 125-131.

ZELNIKER, T. (1971). "Perceptual Attenuation of an Irrelevant Auditory Verbal Input as Measured by an Involuntary Verbal Response in a Selective-Attention Task," J. Exp. Psychol. 87, 52-56.

APPENDIX

Experimenter	Number of Subjects	Delay in msec.	Level or Gain of feedback signal	Error and Duration Measures
Lee (1950b)	5	0,40,140,280	unspecified (high enough to produce substantial disturbance	Total duration
Black (1951)	22	0 to 300 in 30 msec. increments	85-90 dB at earphones	Mean of total duration
Fairbanks (1955)	16	0,100,200, 400,800	30 dB above an arbitrary reference	a) total instances of error, b) total duration c) rate of instance of error, d) correct word rate, e) total word rate f) error word rate
Fairbanks and Guttman (1958)	(based on Fairbanks (1955) experiments)			a) classification of errors according to type, b) number of phonemes per error
Chase (1958)	20	0,216	20 dB above the vocal output	Number of times [b] was repeated in 5 sec.
Kozhevnikov and Chistovich (1965)	5	0,100,210, 300,470	115 dB at earphones	a) phonemes per sec., b) duration of repeated syllables and of segments within syllables, c) number of repetitions in any one instance of repetition.

Experimenter	Number of Subjects	Delay in msec.	Level or Gain of feedback signal	Error and Duration Measures
MacKay 1968 a)	13	200,263	45 dB above vocal output	a) Number of repetitions per syllable
b)	13	0,100,150,200, 263,375,524,750	45 dB above vocal output	b) $\frac{\text{Total duration}}{\text{Number of correct syllables}}$
Huggins (1968)	1	0,80-400 in 10 msec increments	35 dB above vocal output (70-75 dB speaking level)	a) total duration b) number of repetitions c) acoustic segment durations
Fletcher and Yates (1971)	50 in experiments I, II and III, 48 in experiment IV.	200	35 dB above vocal output	Number of times a particular sound was repeated in 5 seconds
Sussman and Smith (1971)	4	0,100,200,300, 400	unspecified (high enough to mask bone-conducted feedback)	Extent of a) jaw opening b) jaw velocity during opening and closing c) duration of jaw activity
Robinson (1972)	4	0,200,225,250, 275,300	unspecified (feedback signal was accompanied by "pink" noise masker)	Number of errors

Appendix B

Pilot Study Passage

[Once upon a time], a number of mice called a meeting, to decide upon the best means of ridding themselves of a cat that had killed so many of them. Various plans were talked about, but none of them chosen. A young mouse came forward, and said that a bell should be hung round the tyrant's neck. Then they would all hear the cat coming, and so be able to escape. The suggestion was received with pleasure by all, until a thoughtful old mouse got up and said, "I consider the plan a very clever one. But I should like to know, who is going to put the bell on the cat?" [It is easier to make a suggestion than to carry it out.]

The sentences in brackets served as frame and were not analyzed.

The passage was broken up into the following phrases for purposes of analyses:

Phrase 1. a number of mice called a meeting,

Phrase 2. to decide upon the best means

Phrase 3. of ridding themselves of a cat

Phrase 4. that had killed so many of them.

Phrase 5. Various plans were talked about,

Phrase 6. but none of them chosen.

Phrase 7. A young mouse came forward,

Phrase 8. and said that a bell should be hung round the tyrant's neck.

Phrase 9. Then they would all hear the cat coming, and so be able to escape.

Phrase 10. The suggestion was received with pleasure by all,

Phrase 11. until a thoughtful old mouse got up and said,

Phrase 12. "I consider the plan a very clever one."

Phrase 13. But I should like to know, who is going to put the bell on the cat?"

Appendix C

Screening Passage

We know the shocks are painful; we have tried them on ourselves and we know that they hurt. But it is stressful for the person who does the shocking too. You may have used shock successfully with a hundred kids, but you are still apprehensive about it; you always think that maybe this kid will be the exception; maybe you will hurt him and it won't do any good. But then when you shock him and you see the self-destructive behaviour stop, it is tremendously rewarding.

Appendix D

Main Experiment Passage

[Some time ago, a group of mice called a meeting to decide upon a good means of ridding themselves permanently of that awful killer cat.] Many a plan was talked over, when an insignificant little mouse came forward and remarked that they should tie a bell around the tyrant's neck. Then in future they would hear that crafty feline approach and so be able to escape. The suggestion was received with much pleasure, until a thoughtful mouse said, "I consider the whole plan a clever one. But I should like to know, who will put the bell on the cat?" [It is easier to make a suggestion than to carry it out.]

The sentences in brackets served as frame and were not analyzed.

The passage was broken up into the following phrases for purposes of analysis:

Phrase 1: Many a plan was talked over,

Phrase 2: when an insignificant little mouse came forward

Phrase 3: and remarked that they should tie a bell around the tyrant's neck.

Phrase 4: Then in future they would hear that crafty feline approach

Phrase 5: and so be able to escape.

Phrase 6: The suggestion was received with much pleasure,

Phrase 7: until a thoughtful mouse said,

Phrase 8: "I consider the whole plan a clever one.

Phrase 9: But I should like to know, who will put the bell on the cat?"

Appendix E

Main Experiment Sentences

<u>Sentence 10</u> /ɔ/

"[In an average theater actors are working on a scene just like the following:]
Enthralled Maud's cautious daughter pauses, for the chalk scrawl on the wall talks of awesome hawk's claws.
[It is quite clear that such scenes must be carefully executed.]"

<u>Sentence 11</u> /i/

"[In an average theater actors are working on a scene just like the following:]
Seated in between leafy trees the keen East Greenville people see the three greedy teams each retreat speedily.
[It is quite clear that such scenes must be carefully executed.]"

<u>Sentence 12</u> /a/

"[In an average theater actors are working on a scene just like the following:]
At half past one the calm father passes gas masks up the shaft after the castle staff rushed up the path.
[It is quite clear that such scenes must be carefully executed.]"

<u>Sentence 13</u> "[In an average theater actors are working on
/u/ a scene just like the following:]
Soon the two fools who choose crude tools lose to the smooth snoops who use goof-proof rules to shoot snooker pool.
[It is quite clear that such scenes must be carefully executed.]"

The sentences in brackets served as frame and were not analyzed.

Appendix F

Symbols used in the Study

Consonants

	Bilabial	Labio-Dental	Dental and Alveolar	Palato-Alveolar	Pala-tal	Velar	Glottal
Plosives	p,b		t,d			k,g	ʔ
Nasals	m		n			ŋ	
Laterals			l				
Fricatives	β	f,v	θ,ð/s,z	ʃ,ʒ			h
Frictionless Continuants and Semi-vowels	w		ɹ,r		j		

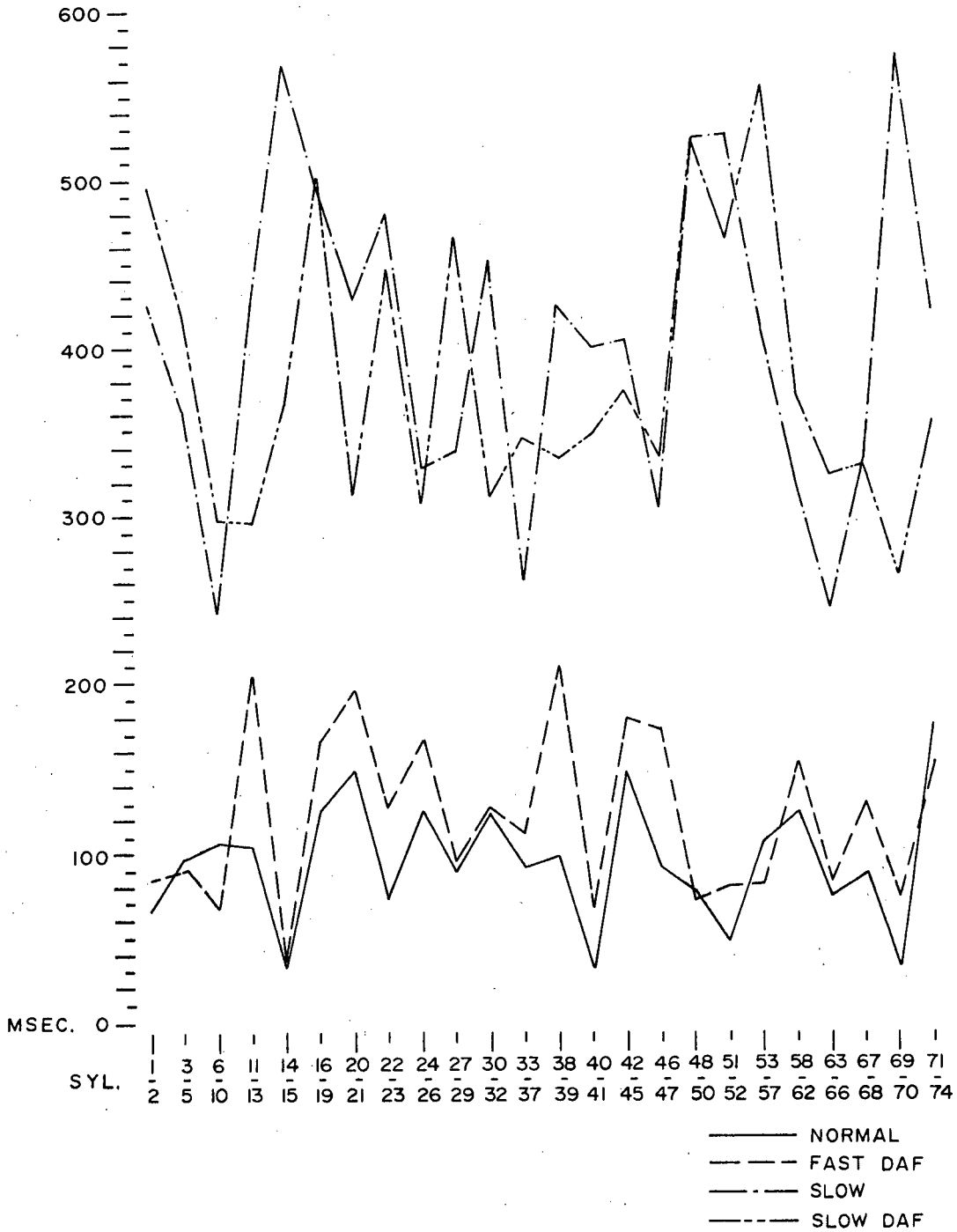
Vowels

	front				back
close	i				u
	ɩ				ʊ
	e				o
		ə			
	ɛ		ʌ		ɔ
	æ				
				ɒ	
open		a			

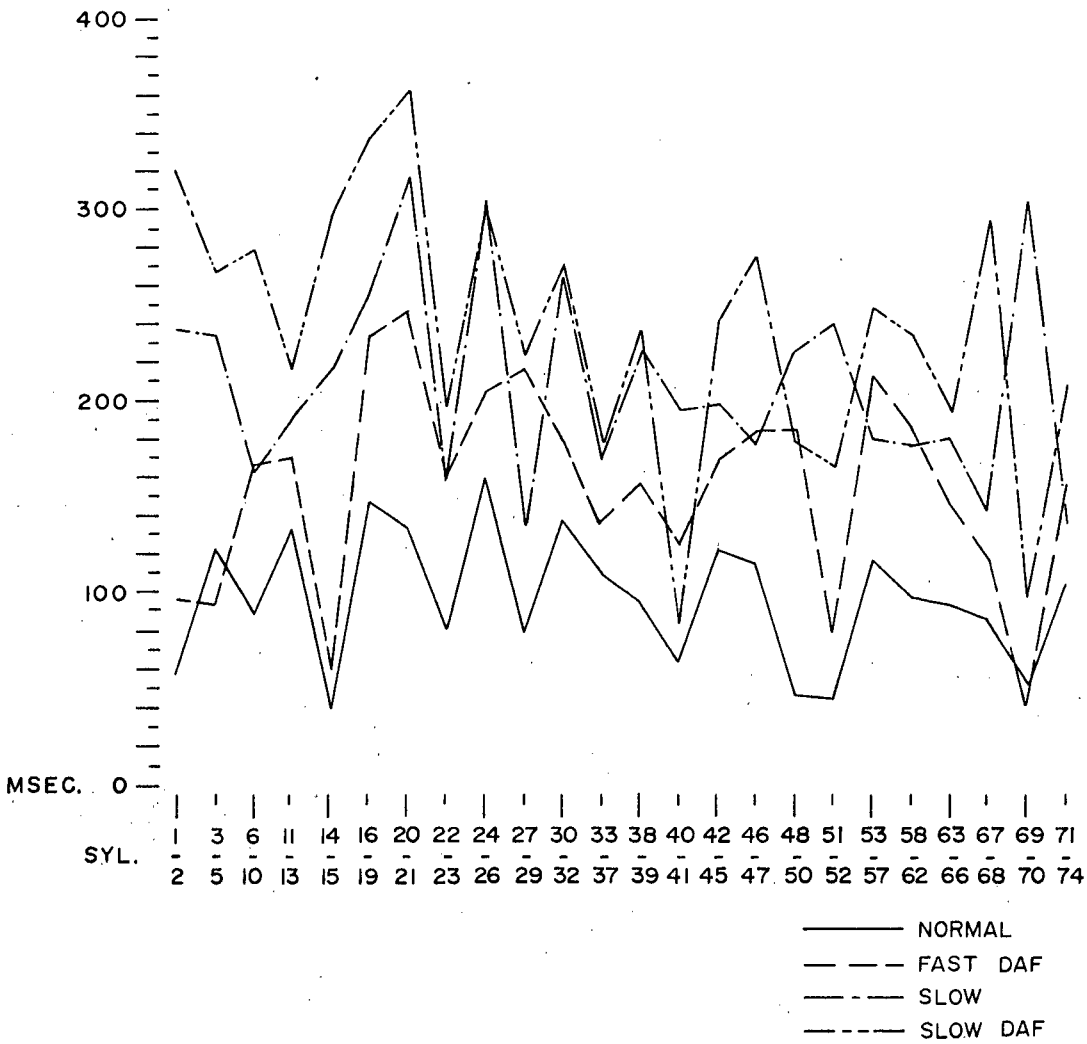
Diacritical Marks:

˺ = unreleased
ˇ = voicing
̥ = breath
̩ = syllabic
ˈ = stress
ʰ = aspiration

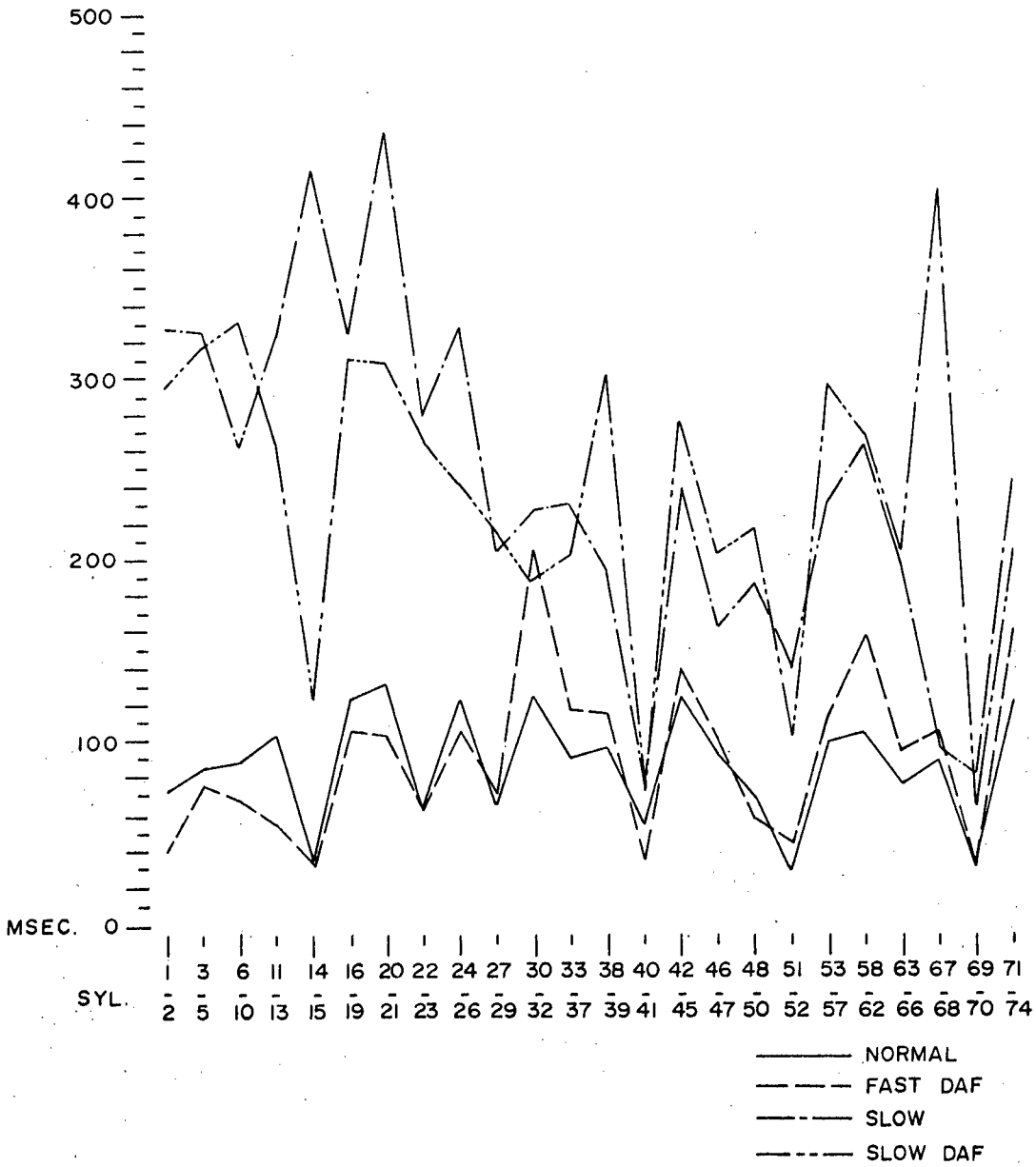
Appendix G



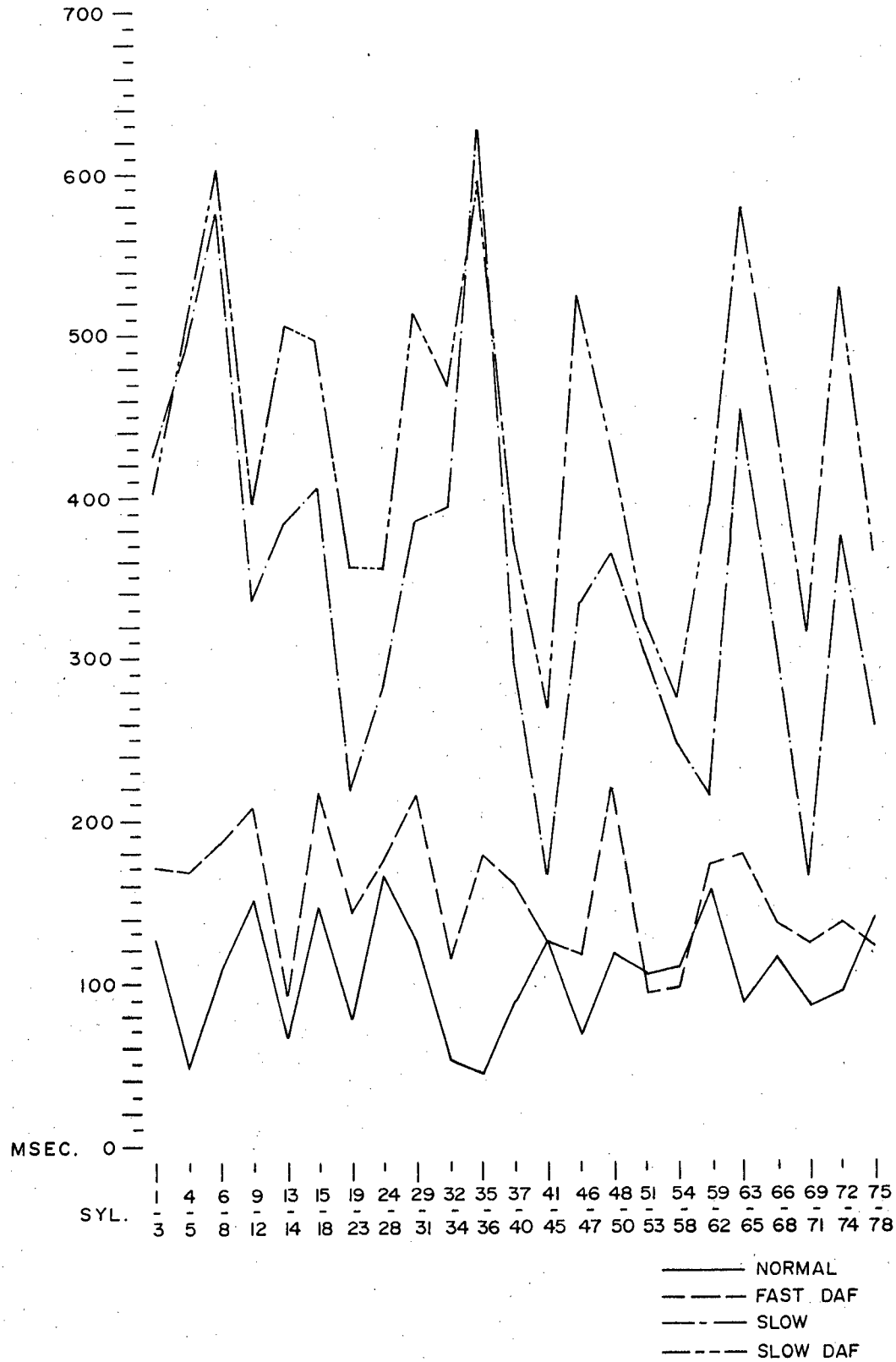
Subject 1, Sentence 12



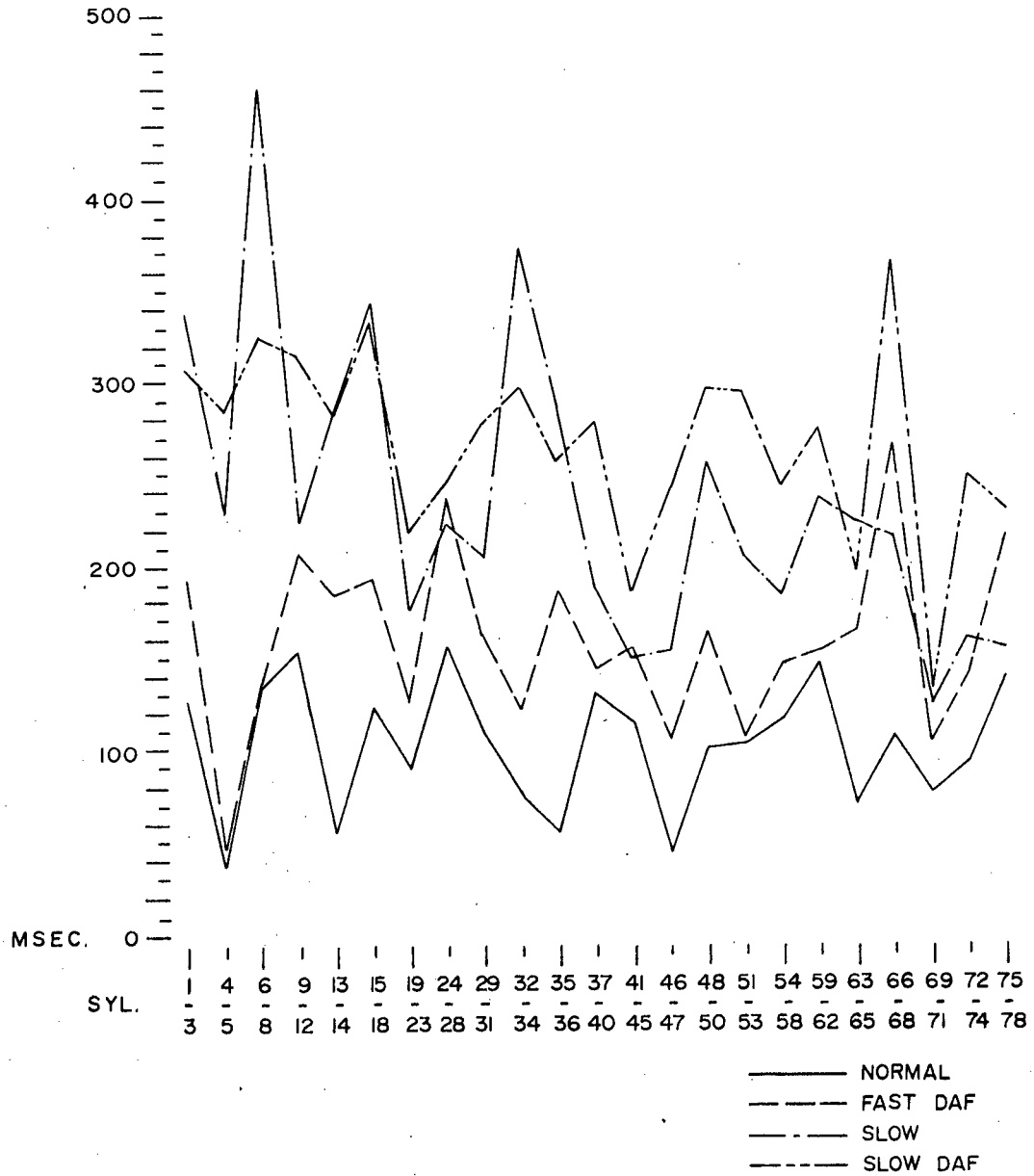
Subject 2, Sentence 12



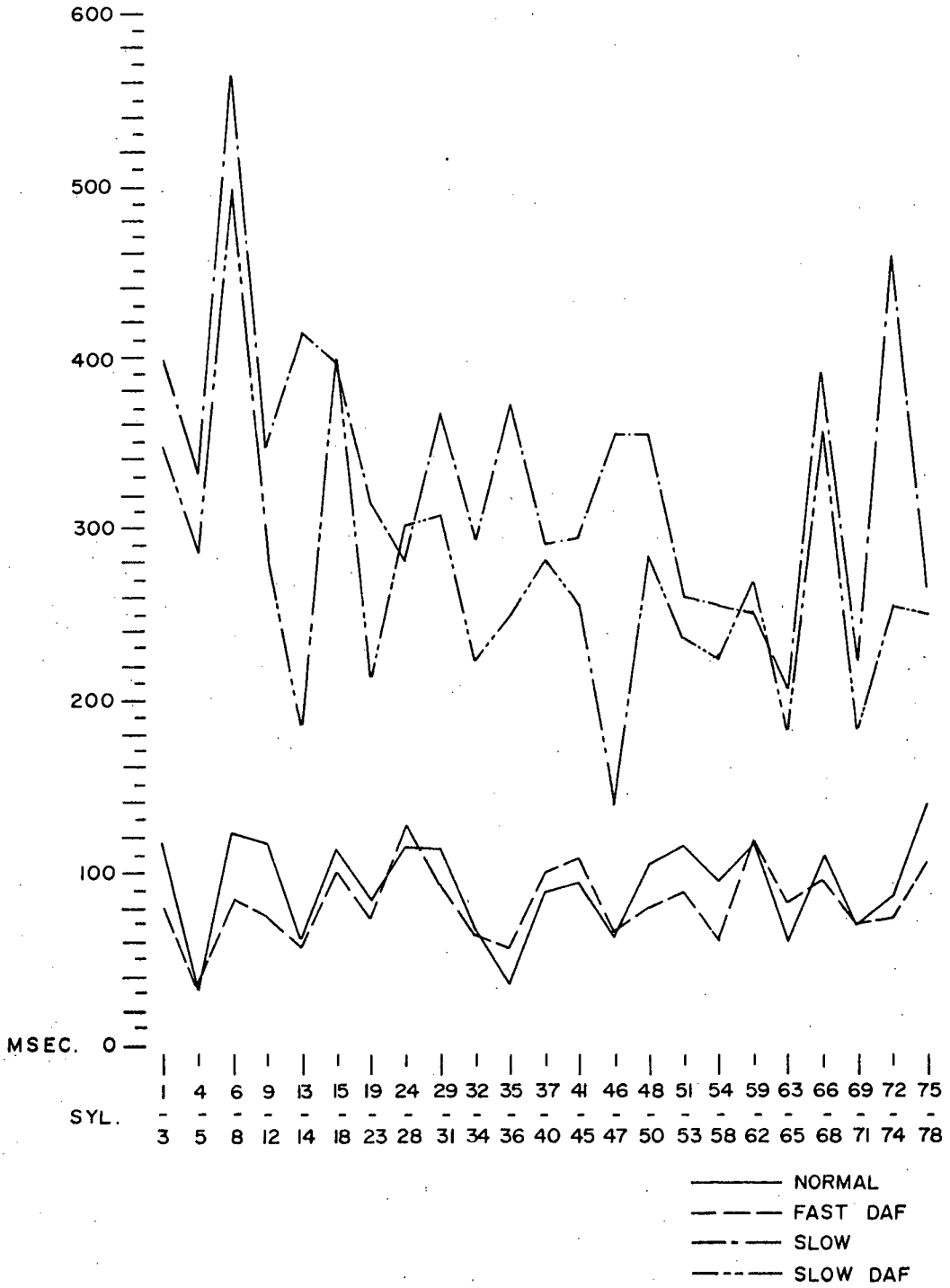
Subject 3, Sentence 12



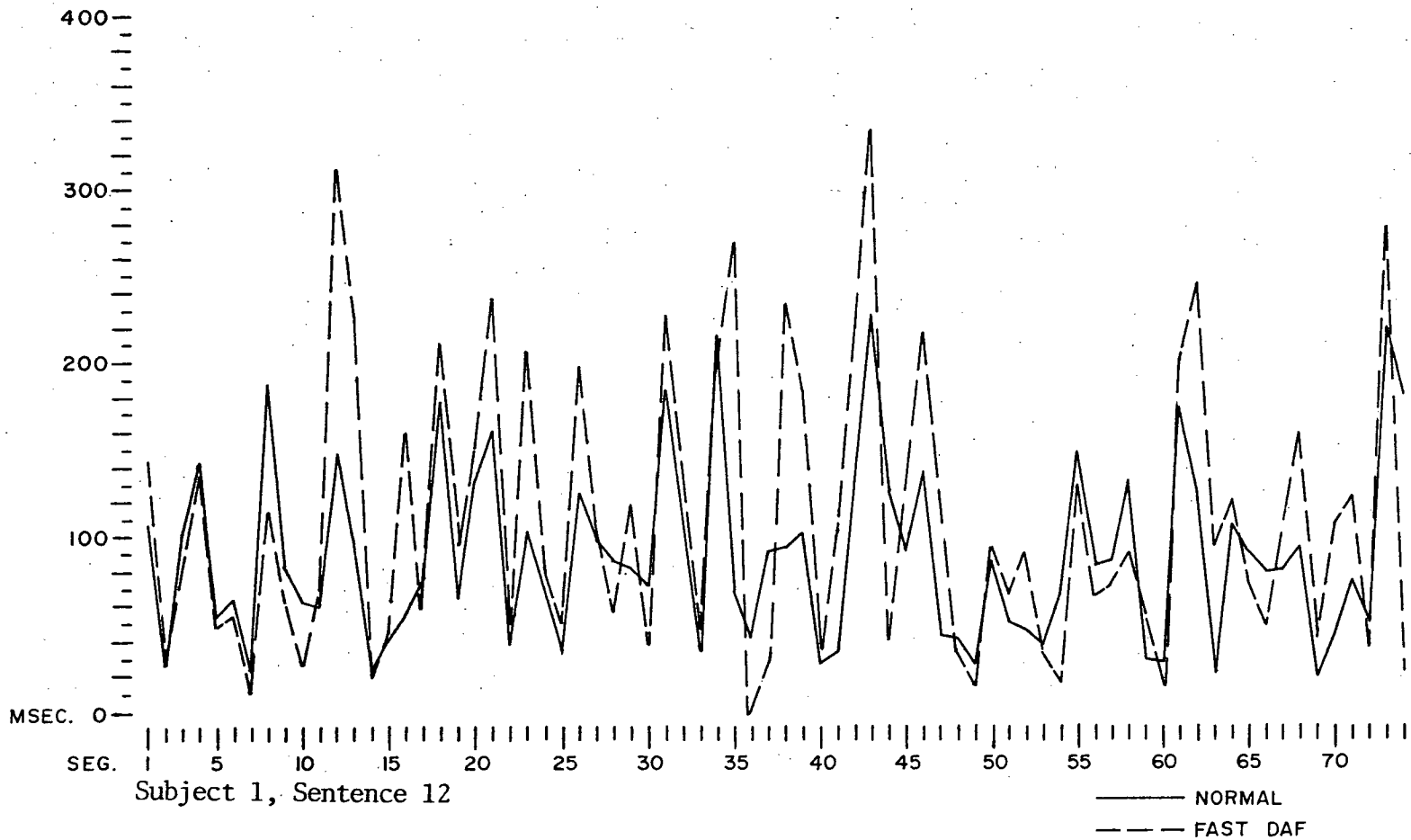
Subject 1, Sentence 13



Subject 2, Sentence 13



Subject 3, Sentence 13



Subject 1, Sentence 12