DURATIONAL CHARACTERISTICS OF SPONTANEOUS SPEECH: EFFECTS OF SAMPLE SIZE ON TEST-RETEST RELIABILITY

by

JANICE USTRUD THOMS

B. A., Augustana College, 1964

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

Approved by:

Bruce C. Flanagan

Major Professor

ACKNOWLEDGMENTS

The following people are to be gratefully acknowledged for their assistance in the preparation of this thesis; Doctor Bruce C. Flanagan, Patricia H. Flanagan and Nancy B. Carlson. The author expresses her gratitude especially to her husband Jerry for his support throughout the project.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

This is a report of a study of the durational characteristics of speech and silence events which occur during spontaneous speaking. The purpose of this research was to ascertain the effects of speech sample length on test-retest reliability of the durational characteristics of speech. Speech may be described as a series of vocal utterances of varying durations separated by intervals of silence which also vary in duration (Minifie, 1963). The present study reports the characteristics of the speech and silence events of 10 male college students. The subjects' spontaneous speech sample was emitted in the presence of pictures and recorded in a sound treated room during two identical procedures on separate days. The length of this spontaneous speech sample was treated as the independent variable, and the difference between measures of speech and silence events between trials was treated as the dependent variable. The purpose of this operation was to establish the relationship between speech sample length and test-retest reliability.

Speech can be treated as a series of sound interspersed with silence events. Implicit in this treatment, however, is the necessity for understanding the complex durational sequences occurring during these speech and silence events. Few speakers are able to organize and verbally communicate their exact thoughts without reiteration or hesitation, and this is constantly manifested and accepted as part of the process of emitting speech. However, the existence of the profession of Speech Pathology, with its therapeutic function, is evidence that the individual's verbal skills are judged and categorized according to how well he approximates certain norms, be they cultural, learned or innate. Several investigators (Flanagan, 1966; Minifie, 1963) postulate that it is perhaps the durational characteristics of one's speech that define what sounds "normal or abnormal" to the listener's ear. Such speech characteristic phenomena described as stuttering, cluttering and sequencing have helped prompt the study of those characteristics of speech somewhat arbitrarily coined "standard" to determine if such norms exist. Shifts of emphasis over the years of investigation show a relatively early differentiation occurring between fluent and disfluent speech, then rather unsuccessful attempts were made to further delineate and categorize disfluent speech, and more recently attempts have been made to define the qualities of fluent speech and its durational characteristics. These recent attempts include the more definitive results of research done on duration of articulatory phenomena and definition of characteristics within durations. Less studied but not less significant have been the actual durations and time patterns occurring during speaking. Studies concerning these phenomena will be treated more extensively in the literature review section of this paper, however one study deserves consideration now because of its significance for the present research. Flanagan (1966) established variance estimates for frequency distributions of speech and silence events for fluent speakers and compared these to frequency distributions of stutterers during a 90 second sample of oral reading. He separated and tabulated speech and silence events according to their durations, assigned these data to arbitrary class intervals, and thereby permitted construction of frequency distributions for speech and silence events. His purpose was to establish whether under certain controlled conditions, the frequency distributions of speech and silence events would demonstrate reliable mean estimates at these class intervals; and whether normal speakers would deviate from nonnormal speakers at at least one of these class intervals. While observing class interval distinctions between stutterers and non-stutterers, the data showed some tendency for fluent individual subjects and subjects as a group to replicate the temporal properties of speech and silence events from trial to trial. However, the low reliability coefficients of the frequency distributions prompted Flanagan to hypothesize that these coefficients were a function of the small sample size of 90 seconds. Flanagan suggested that the next step in pursuit of definitive findings concerning durational characteristics would be to investigate reliability coefficients as a function of the number of responses required from each speaker. Various sizes of speech samples could be analyzed to determine the extent of speech behavior needed to be observed to find reliability coefficients on a test-retest basis. The purpose of this present measurement, then, is to establish whether test-retest reliability of speech and silence events during spontaneous speaking occurs within a sample length of 5 minutes.

CHAPTER II

REVIEW OF LITERATURE

An overview of recent literature about the durational characteristics of normal spontaneous speech shows a concern for the implications of observed time dimensions. These studies can be broadly categorized into four groups:

1. Those studies concerned with gross characteristics of time, that is speaking (and reading) rates and the proportion of total speaking (or reading) time during which phonation occurs;

2. Those studies concerned with articulatory durational phenomena; studies of variations in time denoting distinctive features of phonemes or phonetic variations, and studies defining hesitation, pausal and junctural phenomena;

3. Those studies concerned with recording and analyzing observable time patterns and prescribing models for those time patterns;

4. Those studies concerned with paralinguistic variables[1] which influence time patterns or durational characteristics of speech and silence events.

Such categorization is useful only for organizational purposes and does not suggest 4-point exclusiveness. These areas of research are mutally dependent and often overlapping; and the author's attempt here will be only to cite major studies in order to create a coherent perspective.

---

[1]Paralinguistic is a term used for convenience to embrace such variables as content, emotional state, sex, and situational context. These may ultimately be determined to be linguistic features when their taxonomy has been established.

Studies Concerned with Gross Measurements

One of the earliest techniques used to measure speech and silence events can be described as rate of speaking in words per minute. This is a gross measure used to describe the time dimension of the speaking activity. It is influenced by the durational characteristics of the pauses in speech as well as the speed with which sounds are produced during uninterrupted utterance (Minifie, 1963). This technique was used by Johnson, Darley and Spriestersbach (1963) to establish norms of speaking rate in words per minute. Their work represented the performances of 50 male stutterers, 50 female stutterers and the same number of male and female non-stutterers, respectively, on two speaking tasks. One of the tasks was to talk for 3 minutes about a preferred future job and the other involved 3 minutes of speech about card No. 10 from the Thematic Apperception Test. Computation of words per minute was accomplished by counting only those words spoken without disfluencies, dividing this number by the total number of seconds taken in task completion and multiplying this result by 60. The norms from these data were published and are commonly used by other researchers for data comparison.[2]

Several studies have been done measuring rate of speaking within phrases. Kelly and Steer (1949) recorded the extempore speech of 24 college speakers enrolled in an elementary public speaking course and found extemporaneous speech to be extremely variable in rate. Rate of speaking within

---

[2]Kelly and Steer (1949) purport that overall rate measures do not reveal the variability of speaking and showed that the sentence rate method which they used in studying speakers related more highly to audience judgment in evaluating the speed of these speakers than did over-all rate.

phrases ranged from 125 to 328 words per minute with an average over-all rate of 159 words per minute and an average sentence rate of 209 words per minute. The mean duration of syllables found by Kelly and Steer was .154 seconds, and a correlation of -.47 between mean syllable duration and mean sentence rate in words per minute indicated that the faster the mean sentence rate the shorter the mean syllable duration.[3]

Fonagy and Magdics (1960) found an average speed of utterance (number of sounds uttered in a second) of 11.35 sounds per second in Hungarian when measuring the length of phrases in various speaking and reading contexts of an unspecified number of subjects. An interesting suggestion made was that speed of phrases depended upon their length, which the authors felt was governed by a "tendency to equalization." Shorter phrases were seen to be uttered more slowly than long, although speed dependence was observed to decrease somewhat in phrases of two or three syllables.

A final gross measurement, other than over-all speaking rates and rates within sentences and phrases, is the measurement of proportion of total speaking time during which phonation occurs, or the phonation time ratio. This involves measuring only time spent in speech events during the total time of speaking task. Kelly and Steer (1949) also measured proportion of vocalization time and found their 24 subjects to spend 71 percent of the entire speaking time in phonation. Thus by measuring the proportion of total speaking time during which phonation occurs (phonation time ratio) certain variations in durational aspects of speech can be described which may not be

---

[3]Johnson et al. (1963) point out that -.47 correlation, while greater than chance, suggests that mean sentence rate should not be accounted for wholly on the basis of mean syllable duration.

revealed by changes in over-all speaking rate. For example, it would be conceivable that a person who speaks very rapidly within phrases but takes many or long pauses could have the same over-all rate in words per minute as does one who speaks slowly but with few or short pauses.

Studies Concerned with Micro Measurements

The previous studies have dealt with measurements of relatively large segments of time dimensions of speech. The second category deals with those studies of micro measurements of time factors associated with articulatory phenomena and the analysis of vowels, consonants, and non-speech events. Studies show that in general the duration of vowels preceding voiced consonants is longer than the duration of vowels preceding voiceless consonants. Peterson and Lehiste (1960) extensively studied the influence of preceding and following consonants on the duration of stressed vowels and diphthongs in monosyllabic words, where a large group of words was read by one speaker and a smaller group was read by five speakers. Initial consonant influence upon vowel duration appeared to be negligible. However, the duration ratio of vowels before voiceless and voiced consonants was found to be 2:3, demonstrating that vowel duration is shorter when followed by a voiceless consonant and longer when followed by a voiced consonant. As a class, plosives were found to be preceded by the shortest vowel durations, durations of vowels preceding nazals were approximately the same as those preceding voiced plosives, and voiced fricatives were preceded by the longest vowel durations (Peterson & Lehiste, 1960).

Sharf (1960) measured the duration of vowels before voiced and voiceless stop consonants and found close parallels to the Peterson and Lehiste study. Average duration for vowels preceding stops in monosyllabic words was 222 milliseconds, ranging from 150 to 275 milliseconds as compared to a duration range of 155 to 340 milliseconds found by Peterson and Lehiste. Average duration of vowels preceding consonants in bi-syllabic words was 139 milliseconds, ranging from 92 to 217 milliseconds. As in the Peterson and Lehiste study, a vowel duration ratio of 2:3 was found before voiced and voiceless stops in both monosyllabic and bi-syllabic words. Closure duration of post-stress intervocalic stops was also measured and average duration was found to range from 26 milliseconds for /t/ to 54 milliseconds for /g/. Sharf's findings again show that before voiceless stops, vowels are generally shorter than before voiced stops, although his study suggests that the durational differences in vowels before /t/ and /d/ contribute little to the voice-voiceless discrimination of these two consonants. The reported difference in vowel duration before the voiced and voiceless bilabial stops and the voiced-voiceless velar stops was 30 milliseconds but between the voiced and voiceless alveolar stops the difference was only 9 milliseconds.

House (1961) measured the durations of 12 vowels in bi-syllabic nonsense utterances when spoken by three male subjects. Analysis showed a consistent and significant variation of vowel duration as a function of the phonetic environment for all subjects; the interpretation here being that the voicing characteristic of the consonant influenced the preceding vowel duration. His reported duration range was from 70 to 410 milliseconds for vowels in various phonetic contexts, from 5 to 20 milliseconds for the aspirated portions of voiced stops, and from 50 to 80 milliseconds for voiceless stops.

Huggins (1968) hypothesized that the preferred duration of the /p/-closure, that is the duration perceived as normal, was affected by the following vowel duration. A master sentence was used from which two variations were constructed changing only the stressed vowel in one word: /paupers/. The vowel was shortened by 16 milliseconds for one version and lengthened by 32 milliseconds in the second version. A continuum then was made from each version by altering the duration of the closure of the initial /p/ of /paupers/ in steps of 8 milliseconds (equalling the glottal period in the following vowel). Stimuli on the two continua could then be compared in two ways; either because the duration of the /p/-closure was the same or because the duration of the syllable containing both the /p/-closure and the vowel was the same. The author hypothesized that a change in vowel duration would produce a compensatory change in preferred duration of the /p/-closure. Four subjects judged whether timing in /for paupers/ was lengthened, shortened or not changed. Subject judgments failed to show the expected compensation, but the author felt that the choice of phonemes in the study was partially responsible. Of special implication for patterning was Huggins' observation that subjects did report an increase in speaking rate in the two unstressed syllables preceding a grossly shortened /p/-closure and suggested that the temporal pattern of articulation appeared to be spontaneously restructured to assimulate the shortened closure.

As previously alluded to, studies of durational time variation include attempts to measure, to define and to ascertain the influence of certain features of silence and non-speaking intervals. Liberman, Harris, Eimas, Lisker and Bastian (1961) were interested in the effect of learning on voiced-voiceless phonemic distinction where the experimental variable was duration of silence. Lisker (1957) found that a sufficient cue, although perhaps not the most important, in distinguishing between /b/ and /p/ in intervocalic position was the duration of the silent interval between the first and second syllable of the word. This finding prompted Liberman et al. to further observe the influence of this duration variable. The authors studied the effect of durations of silence intervals from 20 to 120 milliseconds on listener perception of a voiced-voiceless distinction, using both a synthetic speech pattern and a non-speech context. Liberman's synthetic speech pattern consisted of various durations of silence separating the two syllables of the synthesized word. When the duration of the silence between syllables was short the word was heard as /rabid/ and when the silence interval was long was heard as /rapid/. The same durations of silence separated two noise bursts which matched the onset, duration and offset characteristics of the speech signals in the synthesized word. Results of the judgments of 12 university students during the synthetic speech pattern appeared to be more acute across the /p/ /b/ boundary, which lay at about 70 milliseconds of silent interval. In other words, the listeners heard these sounds only as phonemes and did not discriminate any other difference among them. In the non-speech sound task, no increase of discrimination was observed across the phoneme boundary, and discrimination within each pattern was poorer than that observed within the speech sound task. Results indicate, then, a lack of generalization across phoneme boundaries to the noise task.

Micro-Measurement of Non-Speech Durations Associated with Pause. Extensive study has been made of the non-speech phenomena labeled juncture, pause and hesitation. The study by Maclay and Osgood (1959) provides a concise examination of research on silence and pause up to 1959 and treats hesitation phenomena extensively. To aid the reader, the study is here cited for purposes of definition of terms. The authors make a basic distinction between pause and hesitation by defining hesitation pauses as being relatively gross, easily observable and often interruptive to the flow of speech; while juncture pauses are quite short in duration, harder to observe and record, and do not interrupt the flow of speech. Juncture is generally described as functioning syntactically at morpheme, word, phrase and sentence boundaries. The study showed that one-half of all pauses occur within phrases rather than at these boundaries. Hesitation pauses are further described in four ways, "Repetition, False Start, Filled Pause and Unfilled Pause." In studying the locations of these hesitations in speech, the authors found durational influence on Filled and Unfilled Pauses and suggested that the speaker has learned that unfilled intervals of certain lengths indicate a loss of control of the conversation, hence the filled pause, or "ah" is a reaction to his own silence. Lallgee and Cook (1968) caution that this suggestion should be viewed only within the context of the monologue situation studied by Maclay and Osgood.

Lehiste (1960) concisely reviewed the controversy of pause, particularly juncture, beginning with a description of it strictly in terms of modifications of surrounding phonemes (allophonic) and culminating in more recent considerations of its definition in terms of differences in timing or duration. After studying internal open juncture occurring at morpheme and word boundaries, Lehiste concluded that it is a point in time, not a phoneme in its own right and that it serves as a boundary between two bounded sequences.[4]

Boomer and Dittman (1962) differentiate psycholinguistically and durationally between juncture pauses and hesitation pauses. Juncture pauses were found to be under the conscious control of speakers while hesitation pauses were unconscious, and were found to be longer, with an average duration from 500 to 1000 milliseconds. The average duration in milliseconds for hesitation pauses was only 200-250, which does not coincide with findings of Maclay and Osgood who reported junctural pauses to be shorter in duration than hesitation pauses. In the Boomer and Dittman study, sentences were paired and arranged to incorporate certain patterns and durations of juncture and hesitation pause, recorded and played to 25 phonetically naive male and female listeners. The sound sequences in each pair of sentences remained the same, and the difference occurred by varying the type of pause between words, either hesitation or juncture. Listeners were instructed to judge whether the sound sequences were the same or different. One might expect that juncture pauses, because of Boomer's definition of their durational range, would discriminate sentence pairs, however, results showed that listeners discriminated the shorter hesitation pauses significantly better than juncture pauses. The implication made, then, is that definition of speech pause in terms of duration alone "violates certain underlying linguistic and psychological realities" (Dittman and Boomer, 1962, p. 219). However, one might argue that perhaps the definition was not correct in the first place.

---

[4]Acoustically a sequence is bounded if it begins with sound features that characterize the beginning of an utterance (build up) and end with sound features that characterize the end of an utterance (decay).

Barik (1968) suggested that, depending on its duration, a pause occurring between two phonemic clauses can be interpreted as consisting of two components: a juncture pause component associated with the preceding or first clause and a hesitation pause associated with the following or second clause. He questioned Boomer's durational criteria of 200 milliseconds for hesitation pause and 500 milliseconds for junctural pause by citing those studies (notably by Goldman-Eisler) concluding that hesitation pauses function for verbal planning whereas the boundary functioning of juncture pauses is for syntatic purposes. The author suggested that if terminal juncture pauses function grammatically, and hesitation pauses are cognitive, indeed the duration of hesitation pauses should be longer. Barik's rather arbitrary durational criterion, then, for both juncture and hesitation phenomena is 200 milliseconds.

This study might be an example of how a strongly held theoretical position can discount findings, or a unique case in science where the deterministic assumptions do not apply. It would seem that definition of types of pauses occurring in written material could be readily done by careful observation of periods, however, their definitions in spontaneous speech seem rather spurious.

Linguistically, from the findings described in this section we may observe that speech and silence intervals differ in the extent to which they are free to vary in duration. Long silence intervals are relatively free to vary; events of intermediate duration, such as vowels, are less free; and very short events, such as those related to articulatory characteristics (i.e. - stop phases of consonants), are least free to change in duration (Minifie, 1963).

We have seen attempts to measure and quantify durational characteristics of articulatory phenomena and we have discovered that patterns do emerge in these elementary phonations. Vowels of short duration precede voiceless consonants and vowels of longer duration precede voiced consonants; duration of silent interval or stop phases of consonants is closely related to the perception of voicing; and durational characteristics of non-speaking intervals help identify these events. Longer words, phrases and sentences are formed by the grouping of these elementary phonations into longer units. If the durational characteristics of these longer units are to be studied and defined it seems logical to look for patterns in time occurring within this phenomena of duration in speech.

Studies Concerned with Temporal Patterns and Models

The temporal patterns of speech may be described in terms of the frequency distributions of the durations of sound bursts and the intervening durations of silence (Jaffe, 1964). The current thought suggests that both of these distributions are approximately exponential, that is, frequency is inversely related to duration. Controversy over the randomness of the distributional pattern is seen in the literature.

In an early study, time patterns were described by Verzeano (1951, p. 197) as "the ensemble of speed and rhythm in the subject's verbal output exclusive of other vocal variables and or word content." His theory was that the duration of the units of speech vary in a manner not solely determined by their past, and therefore the frequency distributions of the durations of speech units would be a random distribution and would conform to a Poisson series.[5] He postulated that the presence of periodicities (or "abnormalities" such as stuttering) in the time patterns of a subject's speech would appear in the frequency distribution of the units as a departure from randomness; or from the Poisson distribution.

In a following article, Verzeano (1952) studied frequency distributions as they approximated a Poisson series but varied the minimum pause duration of 0.5 used to define a unit of speech. He found a departure from randomness at the short and long duration settings, suggesting that this departure at the small delays occurred because smaller units of speech (0.77 seconds) are more uniform, or equal to each other. He also found that the shorter units (average duration of 1.5 seconds as compared to 0-100 milliseconds in the present study) are more numerous than one would expect from Poisson distributions, indicating that in ordinary speech, shorter words and expressions may be more numerous than long ones.

These studies by Verzeano, and particularly their implication for pathology, prompted Roe, Derbyshire and Sedroski (1958) to investigate speech and silence events to determine whether a particular temporal organization characterized normal speech. If so, then any speech pathology could be described as a modification of this pattern (as suggested by Verzeano). The Roe et al. study showed that it was possible to plot a curve of frequency distributions of normal speakers and that stutterers fell along the outer portions of this curve (vertical and horizontal asymptote). The class interval size for measurement was 50 milliseconds.

---

[5]In this context, a Poisson series refers to the probability that any durational unit will end within 1, 2, 3 ... n seconds after it started.

Flanagan (1966) also studied time patterns and their particular significance for differentiating stutterers and non-stutterers. Preliminary research replicating the aforementioned study (Roe et al., 1958) prompted Flanagan to study the following problems:

1. The establishment of variance estimates for frequency distributions of speech and silence events for fluent speakers during oral reading.
2. The determination of reliability of these frequency distributions on a test-retest basis.
3. Evaluation of the effect of time of day and experimenter variables on these measures.
4. Comparison of frequency distributions of speech and silence between stutterers and non-stutterers.

Subjects for the study were 32 males and 32 females representing fluent speakers, and 10 male and 2 female stutterers. Durational characteristics of the subjects were classified into 14 class intervals of 50 milliseconds each. The important research findings were the differentiation of stutterers and non-stutterers at selected class intervals for both speech and silence events during oral reading. Effects of time of day and sex of experimenter were negligible, and although some tendency was seen for intertrial replication, reliability coefficients were too low to be conclusive.

Minifie (1963) similarily accumulated information concerning the statistical distributions of durational characteristics during speaking and reading tasks, to be used as norms for comparison of deviant speaking behavior. Analysis of the durational characteristics of 12 adult male graduate students was done by an electronic speech duration analyzer which could detect such minute speech and silence events as those occurring during 10 millisecond class intervals. Flanagan (1966) observed that the shape of the cumulative durational frequency curves for silence events in Minifie's study approximated those found by Flanagan's preliminary research (1960). However, Flanagan cautioned against generalization from this data because of the character of a cumulative durational distribution; that is each class interval was dependent on the frequency count which occurred within a fixed time and on events which occurred in every temporal class interval before it. These plots were necessarily employed because of the nature of the measuring device, and therefore subsequent research and instrumentation were proposed which would allow for the collection of data in classes of equal time value; estimations could then be made concerning subject variability at certain points along durational parameters.

Hargreaves (1959) mentioned studies done by Verzeano in which curves showing frequency distribution of speech were J-shaped, that is the shortest speeches were most numerous and the frequency of occurrence regularly decreased as the speech units got longer. He recounted that some authors believe a random process could be a model to approximate distributions and suggested one form of such a model—a version of the exponential decay process. This form assumes that for a person speaking at a point in time, there is some fixed probability that he will stop speaking in the next time interval and that this probability is independent of how long he has been speaking. Hargreaves referred to several of his studies in which speech samples were derived from various situational settings. In each study, distributions of speech unit durations were compared to the model of a simple random decay process and were seen to approximate it in varying degrees. He observed in most cases that the mean utterance duration expressed in seconds seemed to change as a function of situational factors, and therefore proposed that one would have a more accurate model if adjustments for mean utterance duration were made.

Fonagy and Magdics (1960), cited earlier, studied the concept of patterning from the point of view of the "law of equalization" or the idea that the speaker's involuntary endeavor is to pronounce short and long phrases in an approximately equal time. The authors found that this "law" rarely affected phrases of more than six syllables. They further studied the influence of length of phrase on speed of utterance in an unspecified number of subjects by using different types of material. Composite results indicated that a considerable decrease of speed occurred only in the shortest syllables and that the duration of phrases of different lengths was not equalized. Although equalization was not seen phrase by phrase, Fonagy and Magdics stated that speed depends to some extent on duration and can be expressed by an exponential equation.

Another random process suggested as a model for time patterns of speech is the Markovian model. This model, based on the Markov property that the duration of an event is independent of any preceding event, has been studied by Jaffe, Cassota and Feldstein (1964). The authors first generated observed speech and silence frequency distributions based on the Markov property, but these yielded poor approximations to the model. Thus, they assumed that continuation of sound or silence from one unit to the next is more frequent than was predicted from absolute probabilities of the two states, indicating that prior events in the sequence influence subsequent events to some unknown degree. Therefore, a Markov chain with two alternatives (a first order Markov Process) was adopted as the simplest model of dependent probabilities for examination of these sequences.[6] The examiners tested their hypothesis by analyzing 5 minute spontaneous speech monologues from each of 25 subjects. The sampling rate of 200 time units (30 milli-second intervals) per minute yielded 1000 observations for each monologue, and estimations of the number of stops at 200 time units per minute until the matrix achieved a "steady state" indicated a mean value of 4 to 5 seconds. The sequence of the two monologues was randomized and comparisons of obtained frequency means to those predicted showed that a first-order Markov process adequately described time patterns of monologues elicited by the conditions of the experiment. The examiners were also interested in whether monologues prompted by different conditions would be differently approximated by the model, as Hargreaves speculated concerning his model. After experimentation, Jaffe et al. found no evidence to support Hargreave's speculation. In conclusion, the authors conceded to Chomsky (1957), who criticized the Markov process from the standpoint of transformation and further said that the speaker operates simultaneously on lexical and grammatical levels, and the larger and more complex the word and phrase units become the less probable is the sequence. Jaffe stated, however, that although Markovian structures are perhaps inadequate for description of grammatical probability, they seem to describe simple <u>time patterns</u> of speech.

Certain authors suggest that a non-random sequential distribution of pauses exists and contributes to a pattern. Henderson, Goldman-Eisler and

---

[6]The property of this first order process is that the state occupied at time (t+1) is dependent on the state at time, t, but is independent of events prior to t.

Skarbek (1965) measured the distribution of pauses in passages between 350 and 430 continuous words, each passage uttered by one person, and found that there was an alternating sequence with different speech/pause ratios measured in seconds. This study reported that in spontaneous speaking, long pauses and short speech utterances alternated with periods of relatively short pauses and long speech. However, this pattern did not appear in the sample of oral reading, prompting the suggestion that planning occurs during the long hesitation pauses in spontaneous speech and that the alternating patterns observed could represent a cognitive rhythm (here cognitive refers to such processes as formulation of utterances and word finding). The assumption made was that the short utterance-long pause pattern indicated planning and the long speech-short pause pattern was the outcome of the planning period.

Goldman-Eisler (1967) observed the same type of pause and speech distribution and experimented to show the existence of a pattern. Based on a decade of research, the author's assumption was that planning and organization are intrinsic to the production of speech and that hesitation pauses are an index of such cognitive activities. Goldman-Eisler sampled speech of 32 subjects during three speech situations: 10 samples were taken of spontaneous speech, 12 of readings in English, French and German, and 12 of translations of the readings into English, French and German. She found that the longer the total hesitation time during a period of pause, the longer the speech time in the next phrase and stated that pattern relationship is a function of over-all fluency. Results showed that those subjects who spent less than 30 percent of total speaking time in pause time did not show the pattern studied, inferring that if sufficient time was not spent in pause the pattern would not emerge. However, it was also discovered that this alternating rhythm pattern was not a by-product of a 30 percent or longer pause time criterion. Goldman-Eisler observed that 13 out of 32 subjects met pausal time criterion of 30 percent but did not show the pattern. Unlike the Henderson study, this study found an alternating speech-pause pattern in reading, but again only when pause time reached the criterion of 30 percent. Because of the discovery of this pattern in reading, Goldman-Eisler questioned whether temporal pattern is still a manifestation of a cycle of events of planning and verbal production. (This would seem to be in disagreement with Chomsky). Instead, she suggested that patterning is conditioned by degree of hesitancy combined with a complicated feedback process of neurophysiological and tonigenic "tuning". Hesitancy, in Goldman-Eisler's opinion, reflects feedback control and pausing time contains a tonic charge to balance the increase in fluency; this process is then manifested in patterning.

Ramsey (1968) reported patterning findings similar to those of Goldman-Eisler, studying 56 Dutch females and 28 males on the following tasks:

1. One minute of prose reading
2. Three minutes of longer, more difficult prose
3. Three minutes of conversation with examiner
4. Picture description of three TAT cards
5. Stories made up about three other TAT cards

After measuring mean length of utterances, of silences, of units (sound plus subsequent silence) and of sound/silence ratio, he observed that as the complexity of the verbal task increased the silence time increased and speech decreased. He therefore agreed with the assumption of Henderson and Goldman-Eisler that silence is used for higher cognitive activity. However, again contrary to Henderson's findings, Ramsey observed little difference in length of units across the tasks studied. The author concluded that this is a new finding and further proof of the existence of a rhythm or pattern in speech.

Schwartz, Joseph and Jaffe (1968) reviewed the Markovian model in light of the above studies. The authors stated that the consensus of the Henderson and Goldman-Eisler studies is that rhythm in speech appears to be a functional psycholinguistic unit composed of planning and fluency; hence evidence against a Markovian model. To test the assumption of the authors under scrutiny, Schwartz et al. performed a computer similation of sound-silence durations in spontaneous speech, again assuming that these durations are exponentially distributed. They generated a random number and assigned a pause length of Lp, where the mean of the pause distribution was 0.9 seconds. Likewise, a vocalization was generated and assigned a duration of Lv, where the mean of the vocal distribution was 1.85 seconds. With no correlation between successive events, these sound-silence durations were plotted, and regularity, such as seen in the Goldman-Eisler study, was noted. The authors reported that this is due to the fact that in any sequence of random events, runs will occur as appearing to have a structure. Therefore, the observed patterning of speech suggests not cognitive and fluency determinants but rather chance variation in the context of a Markovian model.

Studies have been cited which suggest evidence for speech patterning as a random and a non-random event. Which ever way one interprets the data, the fact remains that time patterning in speech is observed, irrespective of causation or linguistic theory.

The literature cited to this point has concerned measurement and patterning of durational characteristics of speech and has not meant to preclude individual variability in speech. As mentioned earlier, authors have also studied variables which influence these time patterns and durational characteristics of speech. These studies are briefly cited to give coherence and completeness to the literature review, although the reader must bear in mind that units of measurement in these studies are not necessarily the same as the class intervals employed in the present study.

Studies Concerned with Variables and Durational Characteristics

As was previously noted, Ramsey (1968) found little difference in mean duration of units according to speech task, but he did observe variations of durational characteristics of speech and silence to occur in relation to intelligence, personality and sex. His samples' sizes were from 1 to 3 minutes; however, mean duration data were reported in second units.

Bernstein (1962) studied the influence of social class on durational characteristics and reported significant variation between the working class and middle class subjects observed. His sample criterion was a word count of "approximately" 1800 rather than time, making comparison to the data of this present study rather cumbersome.

Goldman-Eisler (1951, 1954a, 1954b, 1956, 1958a, 1958b, 1958c, 1961a, 1961b, 1961c, 1961d, 1961e, 1964, 1965, 1967) also used number of words as her criterion i.e. "5,096 words by 5 patients" (1967) when measuring differences in duration lengths related to various speaking situations. The smallest unit utilized in expressing mean length of duration was 250 milliseconds, still larger than several of the units employed in the present study. Goldman-Eisler also studied rate of speech production, hesitation pauses, speech breathing and patterning, and the interrelationships of all of these under various conditions, particularly during the psychiatric interview. For the sake of brevity, some of the more significant findings in this study are here listed:

1. Variability in total speech time is mainly a function of the time spent in hesitation pauses and not in time spent in articulation.

2. Short utterances are more variable in speech rate than long utterances, and a high degree of individual stability is found for utterances of more than 100 syllables (note disagreement with Verzeano).

3. Interruptions in the speech flow come from both breathing pauses and hesitation pauses, finding that the former vary more with emotionality and the latter with cognitive processes.

4. Hesitation pauses anticipate sudden increases in information or uncertainty in the message being produced.

Webb (1969) studied the effect of two interview situations on subject speech rates of four males and two females utilizing 15 minute sample lengths. During the standardized interview the experimenter's utterances were confined to 5 second durations, and during the non-standard situation experimenter behavior was not standardized. Measurement of durations was again in second units. Results of the study indicated that experimenters' speech rates, per se, influenced subjects' mean and variance of speech rates. Webb also found greater subject homogeneity of variance and shorter pausal durations during the standardized interview than during the non-standardized interview.

Hargreaves (1955) utilized a sample size of 150 speech units when describing each of four speech situations which occurred during one subject's day. Reported results were that speech units were longer when the subject conversed with acquaintances than when he talked to his wife. Here again, mean duration was expressed in seconds.

Starkweather (1959) measured duration in terms of pattern of utterance length in proportion to total number of 1 second class intervals during the speech of 20 subjects. Analysis of spontaneous speech elicited during two role playing incidents revealed no significant differences in group frequency distributions differentiating the two situations.

Flanagan (1966) tested the variables of time of day during which subjects are run and sex of experimenter and found neither variable to have significant influence on durational properties. It is noted that measurement criteria employed by Flanagan closely approximate the class intervals used in the present study.

This reviewed research concerning the durational characteristics of speech and silence events encompasses studies of gross and micro durational measurements, studies of temporal patterning and proposed models, and studies of the influence of certain variables on the durational characteristics of speech and silence. Results of many of these studies remain to be re-evaluated and when possible reconciled. As literature suggests, little research has been directed toward establishing the reliability of the durational properties of speech and silence events of connected speaking (Flanagan, 1966) which is the principle concern of this report.

CHAPTER III

METHOD

Research method of this study is discussed within four subheadings: Selection of Subjects, Apparatus, Recording Procedures and Measurement Procedures.

Selection of Subjects

Ten male college students, ages 18-21 were the subject participants in this research. They were selected from volunteers from the freshman level Oral Communications course at Kansas State University. Aside from availability, subject selection was prompted by several studies. Steer (1936) reported that this college age young adult group, when compared to the entire young adult population, demonstrates greater homogeneity of many gross measures of vocal behavior. Selection of males was in anticipation of possible future comparison of frequency distributions obtained in this research to those obtained from stuttering populations, 9 out of 10 of which are males (Eisenson, 1958).

The 10 subjects were screened by the examiner for both hearing and speech problems and demonstrated no involvement in either area. Subjects were native English speakers but represented assorted familial and ethnocentric backgrounds.

Apparatus

Spontaneous speech of the subjects was detected by an EV665 microphone placed in a sound treated room and recorded by an Ampex AG500-2 which was located in and operated from an adjacent room. Ambient noise level was controlled by recording in a double wall sound treated room. Differences in speech intensity of each subject were adjusted for by regulating the volume of the tape recorder so that subjects' voices peaked the VU meter at zero.

Frequency distributions of speech and silence events were obtained by playing each subject's tape, after rectification and smoothing, to a Hewlett-Packard 132A Oscilloscope and photographing CRO sweeps with an attached polaroid camera. The sweep time of the CRO was set to record at 500 milliseconds per centimeter so that the spot traveling across the scope of 9.4 centimeters registered speech or silence events during 4.7 seconds. A hand switch introduced an external synchronizing voltage which started the sweep and opened the camera shutter and also stopped the sweep and closed the shutter. Between each sweep the spot remained off the view of the screen, the shutter was closed and the polaroid camera advanced along its bed. The hand switch again activated the sweep and the procedure was repeated until 11 sweeps were completed. By interrupting the tape between sweeps it was possible to photographically record a sample of spontaneous speech. Each final photograph showed 11 lines of 4.7 seconds each, sampling sequentially 51.7 seconds of speech. Thus, by moving the camera, interrupting the tape recorder and continuing the process with the hand switch, a series of speech samples was obtained.

To insure that the camera photographed only the speech and silence events of the subject and not extraneous noise, a reference level was set. The setting was crucial, because if the threshold were too low, background noise would generate waves which would be photographed and analyzed as speech. Conversly, if the threshold were set too high, low levels of speech could go undetected and counted as silence. Amplification of the horizontal axis of the CRO spot was set at 20 millivolts a centimeter. The subject's tape was started and the amplification level was then fine tuned so that any audible speech caused the CRO spot to disrupt the horizontal tracing and constant tape noise did not cause deflection. Such fine tuning procedures preceded photographing of each speech sample.

Recording Procedure

Recording procedure will be discussed in terms of room and recording conditions, stimulus material and instructions for vocal responding.

Minifie (1963, p. 13) emphasized the importance of constant recording procedures in minimizing environmental influence on the speech sample, therefore identical room conditions prevailed throughout the research period. The room in which the subject was seated was sound treated and contained the subject's chair opposite the examiner's chair. The only other furniture in the room was a table between the subject and examiner upon which both a typed set of instructions and each picture were placed. A microphone remained at a distance of 15-18 inches from the subject's lips, and the subject was cautioned to avoid obstructing this distance by either the instruction card or the picture, and to keep this distance constant. The recording equipment was operated by a second experimenter in an adjoining room and was also kept constant between research sessions. Standard recording level was maintained by this second experimenter's adjustment of the gain control of the ampex amplifier so that each subject peaked the recording VU meter at zero before speech samples were recorded. The gain control was then left stationary and the recording session was monitored by this second experimenter.

During preliminary study of procedure with five males not included in the experimental study, it was observed that anxiety over possible psychological analysis of what was being said inhibited and thereby shortened responses to the stimulus material. Therefore, before entering the sound treated room with the examiner, each experimental subject was shown the oscilloscope and tape recorder and told that the examiner was studying fluctuations of speech wave forms during spontaneous speech and that his speech would be recorded and then be observed on the scope screen. The subject then entered the sound room with the examiner, was seated, and test procedure began.

Presentation of stimulus material was accompanied by identical instructions read to each subject by the examiner. The 20 pictures from the Thematic Apperception Test designated by the TAT test manual as suitable either universally or for males were selected to be stimulus material for the spontaneous speech sample. Selection of the TAT for stimulus material was based on findings by Johnson, Darley and Spriestersbach (1963) that the TAT is an effective stimulus which increases the probability of speech being emitted. Since the purpose of the research was to observe frequency distributions of durational characteristics as a function of length of sample, explicitly 5 minutes, an adequate sample size needed to be insured. It was recognized that variability in length of response per picture might occur, therefore a minimum time limit for speaking per card was decided upon, but not revealed to the subjects. Using the norms established by Johnson et al. (1963) concerning speaking rate on the TAT task in words per minute—119.2 average for male non-stutterers—and Murry (1943) who reported a 300 word average length of response per card on the TAT; a 2 minute sample length per card was desired. The average number of words per card, then, would ideally be 240, at the lower end of the range reported by Murry. It was assumed that an adequate sample size could be obtained when subjects spoke for 2 minutes in response to each card. These 20 TAT pictures were randomly sequenced and then divided in half; each half randomly alternated among subjects, presenting each subject with 10 pictures during trial I and the other 10 during trial II. To reduce any ordering effect, each set of 10 was again arranged randomly between subjects. This concern with randomizing was to assure that each card would have an equal opportunity to appear at any point in the trial sequence, thereby reducing the probability of obtaining reliability as a function of card content or order. The stimulus material was presented with the following instructions: The first card was placed on the table in front of the subject and the examiner said,

> You will be shown some pictures. Study each picture about 30 seconds and then tell a story about it. You may tell who is in the picture, what happened before the picture, what's happening now and what you think will happen next. Say as much about each picture as you can and try to talk as you ordinarily would.

Subject adaptation to the speaking task was allowed for by presenting the first picture before initiating the recording procedure. While the subject was talking about this first picture, the volume control of the tape recorder was adjusted to the intensity of the subject's speaking level by the second experimenter. Examiner comments and reactions encouraged elaboration if the length of the subject's story did not approximate the predetermined criterion of 2 minutes. These comments were designed to shape the subject's response towards obtaining an adequate sample of his spontaneous speech during presentation of the subsequent cards. Comments were not provided during the recording session because research suggests that experimenter speech rates influence those of subjects (Webb, 1969). The purpose then, of the presentation of card I was to accustom the subject to the recording process, the experimental setting, and to adjust the gain control of the recorder to the intensity of the subject's speaking level. After this preliminary presentation of the first picture, the instruction card was placed before the subject for his reference, if necessary, thereby minimizing examiner interruption of the subject's speech sample thus facilitating recording procedure. Recording began with the examiner telling the subject's name, age and the trial date. The subject then waited 5 seconds and began speaking. The recording procedure followed as previously described.

Measurement

The criterion for measurement used in this research was 100 milliseconds defining a class interval. A speech or silence event was classified according to length of occurrence or duration into the appropriate class interval. Eight 100 millisecond class intervals served as intervals into which each duration of non-production or production was tabulated. These class intervals were of the following durations in milliseconds: x-99, 100-199, 200-299, 300-399, 400-499, 500-599, 600-699, and 700-x. Note that the first and last class intervals are open ended. In the case of the x-99 millisecond interval, the smoothing filter caused a small but undetermined time delay. Known deflection of 30 milliseconds was, however, observable on the scope face. The 700 millisecond and above class interval was previously determined to be asymptotic value with respect to the variables in question (Minifie, 1963; Flanagan, 1966).

The 11 sweeps on each photograph were measured by a metric scaled ruler. The physical measuring process required two people; one to measure the length of events and the other to tabulate the instance of the event in the appropriate class interval on a data sheet. Examiner reliability was established by test-retest procedures for five photographs. Differences of less than two percent were noted between trials. In the process of photographing speech as closely sequential as possible, the shutter was sometimes opened or closed during an event, hence not photographing the beginning or end of it. Those partial events, termed front and end loads, were not considered a part of the data, unless over 700 milliseconds, because either the beginning or the end of the event was not recorded on the picture. To account for these front and end loads in the total time of each sweep, they were measured and then subtracted from the 4.7 second sweep time. This difference then, represented total time of observed speech and silence events per sweep. After computing this actual time of measured durations, the data of trial I and trial II were divided at 30 second units of time and separate sub-totals of silence and speech were taken within each time unit. These sub-totals cumulated at the final unit—300 seconds of speech, or 5 minutes. All additions and totals were performed and cross checked on an adding machine. Subject intertrial reliability was computed by comparing the sub-totals of each time unit at each class interval for speech and silence events, again treating speech and silence separately. Medians of these differences were computed at each class interval for speech and then for silence events.

It is evident that one disadvantage of the described procedure is its lengthiness. However, the advantage of photographing durational characteristics is that one can observe all speech and silence intervals, even those occurring more rapidly than during 50 milliseconds. Should aspects of such analysis prove to be clinically useful, automated measuring devices are available (Flanagan, 1966).

Measurement of rate of speaking and disfluency in words per minute was also done to provide data on the representativeness of the subjects. To determine the rate of speaking, 50-word samples were transcribed from the beginning, middle and end of each tape by the experimenter. Disfluencies were marked but not included in the 50-word count. The length of each taped 50-word sample was timed with a stop watch and an average of the three times obtained was taken. Speaking rate computation was performed on the calculator for each subject during each trial, and represented average speaking rate in words per minute.

Disfluency count was taken from speech samples transcribed, and a disfluency index calculated for each subject. A disfluency was defined as an interjection or repetition of a sound, syllable, word or phrase. The repetition of a word was counted as one disfluency regardless of the number of times that word or a portion of the word was repeated. The same criterion applied to the interjection of a single word or an entire phrase, each being counted as one disfluency. The selection of this procedure was based on research by Sander (1961) and Siegel and Martin (1965) who reported that it yields high reliability coefficients between and within experimenters (above .90). Computation of total disfluency index was accomplished by the calculator.

CHAPTER IV

RESULTS

The purpose of the research was to determine the effect of sample length on test-retest reliability of the durational characteristics of speech and silence events of 10 college age males. Eight arbitrary class intervals ranging from under 100 milliseconds to over 700 milliseconds provided the criteria of measurement for frequency distributions of speech and silence events. Research findings are presented within the context of the following organizational structure:

1. Observed disfluency and speaking rate
2. Intertrial reliability measures for speech events by sample length
3. Intertrial reliability measures for silence events by sample length

Observed Disfluency and Rate of Speaking

The purpose of this measurement was to further substantiate examiner judgment of subject speech normalcy on preliminary speech screening tests. Norms used in subject comparisons of disfluency and rate of speaking were those of Johnson et al. (1963). Reading time was determined by timing three 50 word segments from each taped sample following Johnson's procedures. Results indicate that the data of the present study compare favorably with the established norms for mean rate and range per trial. The present study showed a subject range from 76.8 to 195.6 words per minute on trial I and from 79.8 to 169.8 words per minute on trial II. Mean for speaking time during trial I was 125.9 words per minute and for

TABLE 1

Means and Ranges of Speaking Rate and Disfluency Index of 10 Male Subjects on the TAT Task.

Measure	Thoms	Johnson et al. Norms on TAT
Speaking rate in words per minute		
Trial I:		
Mean Rate	125.9	119.2
Range	76.8–195.6	72.5–197.8
Trial II:		
Mean Rate	125.1	
Range	79.8–169.8	
Disfluency rate in disfluencies per 100 words		
Trial I:		
Mean Rate	9.73	6.6
Range	0–20	0.7–19 .9
Trial II:		
Mean Rate	9.39	
Range	1.3–19 .3	

trial II was 125.1 words per minute showing no shift between trials, and falling within Johnson's range of 72.5 to 197.8 words per minute.

Results of the disfluency index range show a seeming discrepancy; subject mean during both trials was higher than the average figure for Johnson's norms. However, both the norms and the present data were equivalent with respect to range and no shift in subject mean was seen between trials. Mean of disfluencies expressed in disfluencies per 100 words on trial I was 9.73 and on trial II was 9.39, still well within Johnson's range. Disfluency range for the 10 experimental subjects was from 0 to 20 disfluencies per 100 words on trial I and from 1.3 to 19.3 disfluencies per 100 words on trial II. Data on individual subject range show that one subject just failed to approximate the upper limit of the disfluency range on trial I and another did not come within the lower limit of the range, again during trial I.

Intertrial Reliability Measures for Speech Events by Sample Length

The speech events measured were all those vocalizations transcending the intensity level for noise of the oscilloscope. Briefly summarized, frequency distributions for speech events were derived from recorded speech displayed on a cathode ray oscilloscope. The speech events were photographed, measured and tabulated by duration into appropriate class intervals. Speech events in each interval were sub-totaled within 30 second time units and cumulated at 5 minutes for trial I and II separately. Percentage of individual test-retest reliability was measured at each class interval. Medians were then computed by ranking 1 minute individual reliability percentages at each class interval to determine the point

TABLE 2

Medians of Intertrial Reliability at 1 Minute Sample Units for Speech Events of 10 Male Subjects.

Time in Milli-seconds	Median Per Minute				
	1	2	3	4	5
x- 99	72.4	82.4	92.0	91.8	84.4
100-199	57.3	72.2	79.1	75.0	71.5
200-299	77.8	85.2	82.4	75.2	79.0
300-399	77.7	75.6	85.7	82.5	85.5
400-499	70.2	68.3	79.0	84.1	86.7
500-599	73.0	66.7	70.0	81.3	81.3
600-699	75.0	76.0	67.6	74.4	77.0
700-799	58.6	76.4	76.5	73.4	71.4

dividing the data in each minute in half, and appropriate calculations were made.

Table 2 shows medians of intertrial reliability at 1 minute sample units for speech events. A tendency was noted for reliability to increase as sample size increased and then to level off at the point of highest reliability. Data at class intervals of 99 milliseconds and under and of 300 to 399 milliseconds seemed to build up to the highest reliability indication at 3 minutes and then to level off. Class intervals of speech duration beginning at 100 milliseconds and at 700 milliseconds indicated this same tendency but median values were lower. Speech durations of 400 to 499 milliseconds showed a fairly systematic increase in reliability indication, being highest at 5 minutes. Reliability at 500 to 599 milliseconds showed a leveling off effect at 4 minutes. Median values of the class interval beginning with 200 showed a peak at 2 minutes and than a drop at 4 minutes, and medians of the class interval of 600 milliseconds remained fairly constant throughout time.

The speech sample size at which highest reliability was observed for the class interval of under 100 milliseconds was 3 minutes. The class interval of 100 to 199 milliseconds also showed a reliability peak at the 3 minute sample size. Highest reliability for 200 to 299 milliseconds occurred at 2 minutes, while reliability of the class interval of 300 to 399 milliseconds again peaked at 3 minutes. Durations of 400 to 499 milliseconds showed reliability to be highest at the 5 minute sample size. A reliability peak at 4 minutes was observed for speech durations of 500 to 599 milliseconds and the peak for the class interval of 600 to 699 milliseconds occurred at 5 minutes. Reliability high for 700 to 799 milliseconds was observed at 3 minutes. Mean of times at which reliability peaked across intervals was 3.5 minutes, ranging from 77.0 at 5 minutes for the class interval of 600 to 699 milliseconds to 92.0 at 3 minutes for speech events under 100 milliseconds. The modal point for sample size at which highest reliability was noted was 3 minutes. Average reliability of durational characteristics at the highest points was 86 percent.

<u>Intertrial Reliability Measures for Silence Events by Sample Length</u>

Data gathering, measuring and analyzing of silence events were done identically to and simultaneously with those of speech. Table 3 presents medians of intertrial reliability at 1 minute sample units for silence events of 10 male subjects. Again in silence, the tendency was for reliability to peak at a sample length and then to level off. Class intervals of 200, 400 and 600 milliseconds showed a climbing trend culminating in highest reliability points at 5 minutes. Silence events of the first class interval and the 500 millisecond class interval showed median values which peaked at 3 minutes and then leveled off. The silence interval of 100 to 199 milliseconds showed a leveling at 3 minutes, but median value was highest at 2 minutes. This same pattern was seen at the 700 millisecond class interval.

Silence durations of under 100 milliseconds showed highest reliability at the 3 minute sample size. The class interval of 100 to 199 milliseconds showed a reliability peak at 2 minutes. Reliability of class intervals commencing with 200, 300 and 400 milliseconds showed reliability to be highest at the 5 minute sample size. Highest reliability for the class interval of 500 to 599 milliseconds was observed at a 3 minute sample

TABLE 3

Medians of Intertrial Reliability at 1 Minute Sample Units for Silence Events of 10 Male Subjects.

Time in Milli-seconds	Median Per Minute				
	1	2	3	4	5
x– 99	83.6	85.4	88.9	87.5	87.6
100–199	69.8	83.3	79.2	79.4	77.7
200–299	45.0	64.9	74.8	72.8	75.3
300–399	70.9	55.6	60.3	71.8	78.7
400–499	50.6	71.8	71.7	72.7	80.7
500–599	25.0	63.3	74.6	66.6	72.4
600–699	66.7	63.4	64.6	64.6	79.8
700–799	76.4	92.9	87.1	89.1	90.1

length. Reliability of silence durations beginning with 600 milliseconds peaked at 5 minutes and for durations commencing with 700 milliseconds peaked at 2 minutes. The mean time at which these reliability peaks occurred was 3.9 minutes with a range of 75.0 at 5 minutes for the class interval of 200 to 299 milliseconds to 93.0 at 2 minutes for the class interval of 700 milliseconds. The modal point was 5 minutes and the average of reliability peaks of these durational characteristics was 82 percent.

Some consistency in reliability peaks was noted across corresponding class intervals between speech and silence. Class intervals of under 100 milliseconds peaked at 3 minutes, intervals of 400 to 499 milliseconds peaked at 5 minutes and class intervals of 600 to 699 milliseconds also peaked at 5 minutes. Frequency distributions were similar to those reported by Flanagan (1966), and tables presenting individual data are located in the Appendix.

CHAPTER V

DISCUSSION

The research here reported represents an attempt to study, during spontaneous speaking, the test-retest reliability of durational characteristics of speech and silence events as a function of length of sample elicited. Sidman (1960, p. 258), when describing steady state criterion of operant experimentation said, "By following behavior over an extended period of time with no change in the experimental conditions it is possible to make an estimate of the degree of stability that can eventually be maintained." In principle then, findings from the present research suggest just such an estimate of the degree of stability that can be sustained, showing reliability of durational characteristics to occur within a 4 minute speech sample. Durational characteristics of speech events showed a mean reliability of 86 percent, ranging from 77.0 at 5 minutes for the class interval of 600 to 699 milliseconds to 92.0 at 3 minutes for the class interval of under 100 milliseconds. Reliability of silence events averaged 82 percent, ranging from 75.0 at 5 minutes for the durational characteristics beginning with 200 milliseconds to 93.0 at 2 minutes for the class interval of 700 milliseconds and above. Results of the average value of maximum reliability showed a peak at 3.5 minutes for speech events and 3.9 minutes for silence events, indicating that behavior stability may be observed in a 4 minute sample criterion.

Sidman declared that the steady state observed need not necessarily represent the ultimate stabilization point of behavior. He emphasized that the defining characteristic of a steady state criterion was whether it could select a replicable state. Assuming reliability of durational characteristics of speech and silence events to be a function of a 4 minute sample size, sampling speech behavior at 90 seconds (Flanagan, 1966; Ramsey, 1968) could result in fluctuating data mistakenly attributed to certain variables; whereas in reality the behavior observed may have been manipulated before attaining a stable state, and therefore resulting data failed to show orderliness. This same principle is demonstrated in the techniques used by audiologists. Hearing is measured in steps of 5 decibels because testing by this criterion yielded consistent measures of the unadapted level of hearing sensitivity. Testing in steps of 2 decibels did not yield the same reliability but rather showed inter-test fluctuation in subject response. Testing in steps of 10 decibels, while showing reliability, also obliterated lesser changes in hearing sensitivity. The existence of test-retest reliability of the present study at a mean time of 3.7 minutes therefore, suggests that a 4 minute sampling of speech would yield orderly and systematic data.

The present study showed average reliability of speech durational characteristics, 86 percent, to be slightly higher than that for silence, 82 percent. This was consistent with findings of Minifie (1963), Goldman-Eisler (1956), Webb (1969) and others that the durational characteristics of silence events are more variable than durational characteristics of speech. These authors found that variations in speaking and reading rates were more a function of "compression and expansion" (Minifie, 1963, p. 62) of silence intervals than of variations of speech durations. It was observed in the present research that the highest indicators of reliability for four speech intervals, those beginning with x, 100, 300 and 700 milliseconds, occurred uniformly at a 3 minute sample size. As stated in the literature review, vowels range from 70 to 410 milliseconds, encompassing the above cited class intervals, and were found to be relatively limited in the degree to which they are free to vary (House, 1961). The data distributions of silence are more scattered. This trend was also seen in the literature—evidence the rather abortive attempts to assign silence events to fixed durations, culminating in the linguists' statement that durations of silence are most free to vary (Minifie, 1963). Hence, the durations of speech intervals are seen to be more reliable than those of silence.

That this variability in speech and silence is not a singular result of subject selection is evidenced by trial I and II replication of speaking rate and disfluency index. It is recalled that average speaking time during trial I was 125.9 words per minute and for trial II was 125.1 words per minute. The disfluency index indicated a trial I mean of 9.73 as compared to 9.39 disfluencies per 100 words on trial II. Again, demonstration of data reproduction indicates validity of the figures and normalcy of subjects with respect to other measures of durational characteristics and fluency.

It would seem that before experimenting with independent variables as they affect speech, that before studying temporal patterns, or before analyzing frequency distributions of durational characteristics, one must understand the stability of the phenomenon in question. As was explicated in the literature review, little attempt has been made to establish speech behavior stability criterion, yet speech is constantly being manipulated and results of studies published. The literature contains examples of studies analyzing various amounts of speech with little attempt to control this sample length. The question is, what determines for the experimenter, the amount of behavior he studies? How does the reader know whether it is the measurement applied to the behavior or the behavior itself which determines the findings reported by this experimenter? The appearance of test-retest reliability in the present study as a function of a 4 minute sample length may have rather significant implications for these findings of past studies. Henderson et al. (1965), Goldman-Eisler (1967), and Ramsey (1968) in separate studies suggested evidence for non-random patterning in speech yet were refuted when Schwartz, Joseph and Jaffe (1968) demonstrated that by employing the principle of the Markovian model, the same type of alternating pattern was observed as a random phenomenon. However, again within the first three studies mentioned, there seemed only vague concern with sample size and no attempt to control it according to findings on other studies. Henderson studied behavior in approximately 3 minute samples (350-430 words), Goldman-Eisler studied approximately 5 minutes of speech behavior (as calculated from words per minute) and Ramsey studied from 1 to 3 minutes. Recalling the example of hearing measurement, it can be seen that too large a unit can obliterate important data yet too small a unit of measurement may result in fluctuation. It is conceivable therefore, that the observed patterns could have been functions of failure to have the behavior under experimental control. On the other hand, if speech is a non-random event, seemingly more conclusive evidence for non-random behavior would result from employing a reliable sample size.

The same criticism of studies concerning the non-randomness of speech can be applied to those that attempted to define pausal phenomena. Boomer and Dittman (1962) defined the durational threshold of juncture as 500 milliseconds and of hesitation as 200 milliseconds using varying sample sizes. In a seeming contradiction of terms, Maclay and Osgood defined hesitation pauses as being the longer of the two and utilized a sample size of 163 utterances by 13 speakers at a conference, each utterance being not less than 80 words in length, with a mean utterance length of 309 words. Barik (1967), rather diplomatically placed the threshold for both hesitation and juncture phenomena at 200 milliseconds by logical derivation. While this present study does not contribute specific answers to the cause of this controversy, if reliability had been used as the criterion for behavior gathering rather than the rather arbitrarily determined sample sizes, perhaps more meaningful methods of differentiating pauses and hesitations might have emerged.

It will be recalled that Flanagan (1966) noted a tendency for subjects to approximate the durational characteristics of their speech during oral reading samples of 90 seconds from trial I to trial II, but that individual reliability coefficients were low. The data from the present study show lowest reliability indications at 1 minute but steadily increased as sample length increased. Therefore, perhaps the low reliability observed by Flanagan was not a function of the durational characteristics per se, but rather of the length of sample employed, i.e. 1.5 minutes.

Implications for Further Research

Results of the present study indicate that reliability of durational characteristics of speech is observable on a test-retest basis as a function of a 4 minute sample size. However, because of the laborious techniques used in data analysis, direct replication of the present study should not be undertaken lightly. Recommendations for further study of sample size criteria would include the use of automatic duration analyzing equipment. Alternatives for stimulus material might also be developed. Recognizing the difficulty in obtaining samples of conversational speech, it is still felt that the emotionality associated with the TAT may influence sample length desired. Webb (1969) suggested that the Standardized Interview (Matarazzo, 1962) is a valid basic technique for obtaining adequate speech sample length but that Standardized Interview behavior by the experimenter appeared to have definite effects on subjects' speech rates. Therefore it appears that research concerning stimuli effective in obtaining both adequate sample length and representative behavior is needed. It is also suggested that more definitive statements might be made concerning sample size of the present study if inferential statistics could be used as opposed to the descriptive analysis employed. This might imply expanding subject criteria to more than 10. For this study, the problems of obtaining an adequate sample size for inferential statistics were beyond the scope of resources available.

Sidman (1960) suggests that experimental manipulation of steady states would yield data that are orderly and generalizable to other situations. Therefore, other implications for future research might include testing the influence of variables such as listener attitudes, content of stimulus material and history of speaking reinforcement on a controlled sample length. This latter variable might be especially lucrative in view of findings by Lane (1960, 1964) that durational characteristics of the vowel /u/ could be varied according to schedules of reinforcement. Lane reported an increase in mean and variance of the response amplitude, fundamental frequency and duration of /u/ of four humans under the following sequence of reinforcement conditions; CRF (continuous reinforcement), VI (variable interval reinforcement and EXT (extinction). The same changes in vocal behavior parameters were observed in the behavior of a chick under this schedule of reinforcement (Lane, 1964).

As Sidman (1960) suggests, the reconciliation of previous research with existing information is a most satisfactory demonstration of reliability and generality, therefore, parts of Flanagan's research dealing with individual test-retest reliability could be replicated, treating sample length as the independent variable. Establishment of intertrial reliability would be most expedient for the profession of Speech Pathology in its search for more definitive methods of measuring fluency and disfluency. Assuming systematic reliability of durational characteristics can be observed in speech, then perhaps the techniques of Lane could be applied to programming the speech of stutterers according to norms of fluent behavior.

Further research comparing frequency distributions and reliability of durational characteristics of spontaneous speaking with those of oral reading might be undertaken. Flanagan's study, which prompted the hypothesis of this research, was concerned only with the durational characteristics of speech and silence during oral reading. In view of findings concerning the variability of spontaneous speaking (Kelly and Steer, 1949; Goldman-Eisler, 1958; Webb, 1969), reliability indications might not be expected to be as high as those found during oral reading. Minifie (1963) noted significant difference between the durational characteristics observed during oral reading and impromptu speaking and stated that the largest changes which occurred in impromptu speaking were in the durations of silence intervals. In view of this supposition that silence intervals are found to be significantly more variable than speech intervals, and that the data from the present study, while reflecting this variability, also showed uniform trends, it is felt that the proposed sample size of 4 minutes would yield orderly behavior for spontaneous speaking.

CHAPTER VI

SUMMARY

The purpose of this research was to demonstrate test-retest reliability of durational characteristics of speech and silence events as a function of sample length during spontaneous speaking. Ten young adult males showing no speech or hearing pathology were the subjects for the study. The spontaneous speech of these subjects was elicited by the picture Thematic Apperception Test, and recorded on two occasions separated by 24 hours.

Speech was described as periods of silence events interspersed with periods of speech events. Frequency distributions for these speech and silence events were derived from recorded speech displayed on a cathode ray oscilloscope. The speech and silence events were photographed, measured and tabulated by duration into appropriate class intervals. The number of speech events within each interval was sub-totaled within time units of 30 seconds and cumulated at 5 minutes for trial I and then for trial II. Percentage of individual test-retest reliability was measured at each class interval for each 30 second time unit. Median values were then computed by ranking individual percentages of reliability at each class interval for 1 minute units and performing the appropriate computations.

Reliability was seen to peak at a mean time of 3.5 minutes for speech intervals and at a mean time of 3.9 minutes for silence intervals. The tendency was observed for reliability of durations of speech and silence events to increase with sample length increase and then to stabilize after the point of highest reliability was attained. Previous research findings indicating that durations of silence events are more variable than speech were confirmed in this study.

The observed steady states of the durational characteristics in the present study indicate that a 4 minute sample of speech represents an adequate gathering of spontaneous speaking behavior from which to establish frequency distributions for duration of speech and silence events.

APPENDIX

INTERTRIAL RELIABILITY DATA OF INDIVIDUAL SUBJECTS FOR SPEECH AND SILENCE EVENTS BY SAMPLE LENGTH

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 99 Milliseconds and Under for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	62.9	78.4	81.0	89.9	80.9
S- 2	70.5	75.6	88.2	93.6	97.1
S- 3	91.7	88.5	95.9	85.9	74.1
S- 4	55.0	77.1	95.8	96.6	82.1
S- 5	96.3	70.3	67.9	83.0	74.3
S- 6	74.3	95.2	84.5	82.2	84.3
S- 7	77.4	85.2	98.6	96.9	94.4
S- 8	69.6	83.0	96.7	94.3	98.8
S- 9	69.8	88.9	98.9	94.5	84.4
S-10	87.9	81.7	72.9	80.7	85.1

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	69.2	75.6	81.7	86.6	82.5
S- 2	51.4	51.5	57.4	71.9	83.3
S- 3	82.1	92.0	92.7	80.6	74.3
S- 4	89.1	81.6	89.4	92.1	86.3
S- 5	87.5	86.4	88.4	99.2	97.9
S- 6	85.1	94.5	83.0	80.1	85.2
S- 7	95.5	88.1	99.8	99.6	99.3
S- 8	82.2	84.5	91.9	87.4	88.9
S- 9	65.5	77.5	78.5	87.5	98.8
S-10	95.1	92.8	90.2	96.1	98.1

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 100-199 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	40.6	46.0	55.8	68.2	70.6
S- 2	45.5	48.6	46.2	59.7	62.7
S- 3	92.9	90.7	90.8	70.4	55.5
S- 4	100.0	76.2	96.6	83.3	83.0
S- 5	68.2	83.3	96.8	85.5	82.4
S- 6	64.5	97.7	78.3	72.0	85.3
S- 7	44.4	45.0	59.6	78.0	83.6
S- 8	42.9	41.2	53.0	63.1	66.7
S- 9	50.0	68.2	96.9	88.2	69.7
S-10	72.5	82.9	89.9	87.1	90.0

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	25.0	57.9	53.3	56.4	51.0
S- 2	83.3	90.5	79.4	84.4	80.0
S- 3	55.6	81.6	78.9	76.8	87.2
S- 4	80.0	75.0	66.7	64.0	66.0
S- 5	100.0	85.0	88.9	99.2	88.5
S- 6	84.6	92.0	86.5	70.9	75.4
S- 7	70.0	86.7	85.0	93.1	89.5
S- 8	65.3	63.1	69.4	58.2	71.4
S- 9	54.5	89.5	100.0	83.3	75.0
S-10	69.6	67.4	76.7	81.9	90.0

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 200-299 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	60.0	84.2	90.6	92.4	86.6
S- 2	14.3	17.2	28.6	37.9	47.1
S- 3	85.2	88.0	84.0	74.5	69.6
S- 4	78.6	69.7	82.2	80.0	80.3
S- 5	78.9	86.1	82.5	91.4	96.1
S- 6	100.0	93.0	82.6	74.5	73.9
S- 7	76.5	93.3	71.7	75.8	96.1
S- 8	76.9	64.4	68.3	70.0	72.2
S- 9	95.0	88.6	86.0	100.0	78.6
S-10	54.5	64.1	73.1	74.0	79.3

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	0.0	41.7	31.3	76.5	56.0
S- 2	66.7	85.7	72.7	50.0	44.0
S- 3	40.0	71.4	91.7	88.2	96.0
S- 4	50.0	75.0	80.0	100.0	87.5
S- 5	25.0	40.0	35.7	29.4	45.0
S- 6	28.6	31.6	54.5	69.2	78.6
S- 7	71.4	58.3	76.9	65.0	55.6
S- 8	36.4	50.0	53.3	66.7	72.0
S- 9	62.5	91.7	80.0	79.2	88.5
S-10	55.6	80.0	90.0	86.8	86.7

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 300-399 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	87.5	100.0	100.0	98.1	89.1
S- 2	75.0	88.9	87.1	80.4	100.0
S- 3	100.0	86.8	98.1	88.2	100.0
S- 4	63.6	68.3	63.0	71.2	72.7
S- 5	60.0	68.2	60.5	56.4	59.4
S- 6	82.4	92.9	95.7	88.0	81.0
S- 7	62.5	66.7	71.1	79.6	80.0
S- 8	86.7	95.8	97.2	73.7	80.3
S- 9	54.2	59.0	62.2	85.9	98.7
S-10	93.8	74.4	84.3	92.5	97.6

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	0.0	80.0	77.8	50.0	36.0
S- 2	100.0	42.9	66.7	68.8	50.0
S- 3	100.0	55.6	50.0	88.2	94.4
S- 4	100.0	80.0	75.0	80.0	91.7
S- 5	66.7	100.0	100.0	100.0	83.3
S- 6	75.0	50.0	50.0	71.4	82.4
S- 7	100.0	16.7	50.0	50.0	50.0
S- 8	66.7	44.5	75.0	93.3	85.0
S- 9	12.5	55.6	53.8	72.2	75.0
S-10	50.0	55.6	43.8	57.1	65.4

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 400-499 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	90.0	100.0	78.6	64.9	75.0
S- 2	44.5	66.7	73.7	84.0	85.3
S- 3	85.7	68.2	89.3	97.4	100.0
S- 4	50.0	58.8	82.6	93.3	97.3
S- 5	62.5	78.6	100.0	92.6	96.7
S- 6	78.6	90.0	93.5	88.4	90.0
S- 7	100.0	100.0	76.9	80.6	94.9
S- 8	40.0	50.0	74.1	81.8	78.6
S- 9	25.0	66.7	60.6	84.2	86.4
S-10	77.8	68.4	79.3	70.5	83.3

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	66.7	100.0	100.0	72.7	71.4
S- 2	0.0	50.0	60.0	53.3	93.3
S- 3	75.0	100.0	100.0	88.9	91.7
S- 4	0.0	0.0	25.0	66.7	62.5
S- 5	66.7	50.0	44.4	50.0	72.7
S- 6	33.3	80.0	83.3	90.0	92.9
S- 7	42.8	100.0	100.0	81.3	65.0
S- 8	33.3	50.0	54.5	72.7	100.0
S- 9	66.7	63.6	56.3	60.0	73.9
S-10	60.0	83.3	93.3	100.0	87.5

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 500-599 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	37.5	58.3	86.7	90.0	78.6
S- 2	37.5	50.5	66.7	91.7	89.7
S- 3	83.3	66.7	55.6	63.6	78.6
S- 4	87.5	82.4	79.2	83.9	88.9
S- 5	100.0	80.0	73.3	68.4	85.7
S- 6	62.5	62.5	65.0	100.0	90.6
S- 7	85.7	100.0	68.2	69.2	69.0
S- 8	28.6	31.3	65.0	65.2	71.4
S- 9	60.0	66.7	100.0	88.9	83.9
S-10	100.0	78.6	76.2	78.6	71.4

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	0.0	22.2	18.2	28.6	33.3
S- 2	66.7	42.9	62.5	62.5	45.5
S- 3	0.0	80.0	35.6	80.0	64.7
S- 4	66.7	60.0	80.0	80.0	85.7
S- 5	100.0	87.5	69.2	70.6	63.2
S- 6	0.0	100.0	100.0	100.0	80.0
S- 7	0.0	66.7	100.0	62.5	62.5
S- 8	0.0	0.0	25.0	40.0	81.9
S- 9	50.0	50.0	88.9	91.7	80.0
S-10	66.7	100.0	80.0	64.3	73.7

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 600-699 Milliseconds for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	0.0	75.0	62.5	66.7	76.9
S- 2	75.0	57.1	72.7	85.7	65.0
S- 3	85.7	77.8	60.0	68.8	78.9
S- 4	75.0	70.0	100.0	73.7	85.7
S- 5	50.0	37.5	30.0	36.4	33.3
S- 6	75.0	85.7	90.0	76.9	92.9
S- 7	75.0	88.9	91.7	75.0	84.6
S- 8	75.0	83.3	62.5	85.0	95.2
S- 9	100.0	76.9	73.7	84.6	70.6
S-10	0.0	8.3	13.3	18.8	33.3

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	33.3	37.5	62.5	60.0	90.0
S- 2	66.7	60.0	42.9	77.8	90.0
S- 3	80.0	75.0	72.7	83.3	84.6
S- 4	0.0	66.7	66.7	50.0	75.0
S- 5	50.0	40.0	35.7	38.8	40.9
S- 6	33.3	45.5	53.3	62.5	55.0
S- 7	40.0	80.0	100.0	66.7	54.5
S- 8	66.7	50.0	25.0	33.3	40.0
S- 9	66.7	72.7	81.3	88.2	95.0
S-10	75.0	100.0	80.0	90.0	85.7

Individual Intertrial Reliability Indicators at 1 Minute Sample Units for 10 Male Subjects at Class Interval of 700 Milliseconds and above for Speech and Silence

Subject	Reliability Per Minute for Speech				
	1	2	3	4	5
S- 1	50.0	63.6	70.6	59.3	70.9
S- 2	60.0	39.1	46.4	37.8	41.2
S- 3	28.6	85.0	96.4	91.9	69.5
S- 4	33.3	75.0	77.3	72.4	71.8
S- 5	80.0	66.7	58.8	55.0	52.0
S- 6	100.0	88.9	91.7	90.9	78.1
S- 7	100.0	92.3	95.2	93.5	92.1
S- 8	57.1	73.3	76.0	84.4	90.2
S- 9	92.9	92.0	97.2	86.0	72.1
S-10	42.9	77.8	64.7	56.0	56.3

Subject	Reliability Per Minute for Silence				
	1	2	3	4	5
S- 1	70.0	73.2	79.7	84.0	85.4
S- 2	70.4	93.0	86.4	95.2	97.1
S- 3	93.8	93.1	75.0	86.9	86.5
S- 4	77.3	92.7	87.7	88.6	93.4
S- 5	85.7	90.5	92.0	95.1	98.0
S- 6	75.0	63.9	78.8	90.8	87.8
S- 7	80.0	89.2	96.3	98.5	96.4
S- 8	77.3	100.0	90.0	89.7	95.7
S- 9	100.0	95.8	79.5	74.1	80.6
S-10	92.3	95.2	90.0	74.0	73.4

REFERENCES

Abercrombie, D. (Ed.) *In honor of Daniel Jones; papers contributed in the occasion of his eightieth birthday, 12 September 1961.* London: Longmans, 1964. Pp. 474.

Barker, J. L. Syllable and word division in French and English. *Manual of Phonetics*, 19, 321-336.

Bastian, J., Delattre, P., & Liberman, A. M. Silent interval as a cue for the distinction between stops and semi-vowels in medial position. *Journal of the Acoustic Society of America*, 1959, 31, 1568. (Abstract)

Bastian, J., Eimas, P. D., & Liberman, A. M. Identification and discrimination of a phonemic contrast induced by silent interval. *Journal of the Acoustic Society of America*, 1961, 33, 842. (Abstract)

Barik, H. C. On defining juncture pauses: A note on Boomer's hesitation and grammatical encoding. *Language and Speech*, 1968, II, 3.

Bernstein, B. Linguistic codes, hesitation phenomena and intelligence. *Language and Speech*, 1962, 5, 31-46.

Bernstein, B. Social class, linguistic codes and grammatical elements. *Language and Speech*, 1962, 5, 221-240.

Black, J. W. Relationships among fundamental frequency, vocal sound pressure and rate of speaking. *Language and Speech*, 1961, 4, 196-199.

Blankenship, J., & Christian, K. Hesitation phenomena in English speech: A study in distribution. *Word*, 1964, 20(3), 360-371.

Blankenship, J. S. Stuttering in normal speech. *Journal of Speech and Hearing Research*, 7, 95-96.

Bloomfield, L. Language. New York: H. Holt & Co., 1933.

Bolinger, D. L., & Gerstman, L. J. Disjuncture as a cue to constructs. Word, 1957, 13(2), 246-255.

Bolinger, D. L. Length, vowel, juncture. Linguistics, 1963, I, 5-29.

Boomer, D. S., & Dittman, A. T. Hesitation, pauses and juncture pauses in speech. Language and Speech, 1962, 5, 215-220.

Boomer, D. S. Hesitation and grammatical encoding. Language and Speech, 1965, 8, 148-158.

Brown, S. F. The loci of stutterings in the speech sequence. Journal of Speech Disorders, 1945, 10, 181-192.

Brown, R. W., & Hildum, D. C. Expectancy and the perception of syllables. Language, 1956, 32, 411-413.

Chomsky, N. Syntactic structures. 's-Gravenhage: Mouton, 1957.

Chomsky, N., Halle, M., & Lukoff, F. On accent and juncture in English. In M. Halle (Ed.), For Roman Jakobson. The Hague: Mouton & Co., 1956. Pp. 65-80.

Clevenger, T., Jr. Coincidental variation as a source of confusion in the experimental study of rate. Language and Speech, 1963, 6, 144-150.

Cotton, J. C. A new concept of speech rate variation. Speech Monographs, 1936, 3, 112.

Cowan, J. M., & Bloch, B. An experimental study of pause in English grammar. American Speech, 1948, 23, 89-99.

Dittman, A. T., & Wynne, L. C. Linguistic techniques and the analysis of emotionality in interviews. Journal of Abnormal and Social Psychology, 1961, 63, 201.

Eisenson, J. Stuttering; a symposium. New York: Harper, 1958.

Eliason, N. E. On syllable division in phonemics. Language, 1942, 18, 144-147.

Fay, W. H. Temporal sequence in the perception of speech. The Hague: Mouton & Co., 1966.

Feldstein, S., & Roglaski, C. Predictability and disruption of spontaneous speech. Language and Speech, 1966, 9, 137-152.

Flanagan, B. C. An analysis of the durational aspects of connected speech with reference to stuttering. Unpublished doctoral dissertation, University of Florida, 1966.

Flanagan, B. The distribution of pause durations in speech. Southern Illinois University, TR-4, Sept. 1960, on AFCRC Contrad AF 19(604) - 6127.

Fonagy, I., & Magdics, K. Speed of utterance in phrases of different lengths. Language and Speech, 1960, 3, 179-192.

Gleason, H. A., Jr. An introduction to descriptive linguistics. New York: Holt, Rinehart & Winston. Rev. ed., 1961.

Goldman-Eisler, F. The measurement of time sequences in conversational behavior. British Journal of Psychology, 1951, 42, 355.

Goldman-Eisler, F. On the variability of the speed of talking and on its relation to the length of utterances in conversations. British Journal of Psychology, 1954, 45, 94-107. (a)

Goldman-Eisler, F. A study of individual differences and of interaction in the behavior of some aspects of language in interviews. Journal of Mental Science, 1954, 100, 177-197. (b)

Goldman-Eisler, F. The determinants of the rate of speech output and their mutual relations. Journal of Psychosomatic Research, 1956, 1, 137.

Goldman-Eisler, F. The predictability of words in context and the length of pauses in speech. Language and Speech, 1958, 1, 226-231. (a)

Goldman-Eisler, F. Speech production and the predictability of words in context. Quarterly Journal of Experimental Psychology, 1958, 10, 96. (b)

Goldman-Eisler, F. Speech production and predictability of words in context and the length of pauses in speech. Language and Speech, 1958, 1, 96. (c)

Goldman-Eisler, F. A comparative study of two hesitation phenomena. Language and Speech, 1961, 4, 18-26. (a)

Goldman-Eisler, F. Continuity of speech utterance, its determinants and its significance. Language and Speech, 1961, 4, 220-231. (b)

Goldman-Eisler, F. The distribution of pause durations in speech. Language and Speech, 1961, 4, 232-237. (c)

Goldman-Eisler, F. Hesitation and information in speech. In C. Cherry (Ed.), Information Theory, Proceedings of the 4th London Symposium of Information Theory. London: Butterworth, 1961. Pp. 176. (d)

Goldman-Eisler, F. The significance of changes in the rate of articulation. Language and Speech, 1961, 4, 171-174. (e)

Goldman-Eisler, F. Sequential temporal patterns and cognitive processes in speech. Language and Speech, 1967, 10, 122-132.

Goldman-Eisler, F., Skarbek, A., & Henderson, A. Cognitive and neuro-chemical determination of sentence structure. Language and Speech, 1965, 4, 18.

Halle, M. (Ed.) For Roman Jakobson. Essays on the occasion of his 60th birthday, 11 October 1956. The Hague: Mouton & Co., 1956.

Hanah, E. P., & Engler, L. Juncture phenomena and the segmentation of a linguistic corpus. Language and Speech, 1967, 10, 228-234.

Hargreaves, W. A., & Starkweather, J. A. Collection of temporal data with the duration tabulator. Journal of the Experimental Analysis of Behavior, 1959, 2, 179.

Hargreaves, W. A. A model for speech unit duration. Language and Speech, 1960, 3, 164-173.

Harris, K. S., Bastian, J., & Liberman, A. M. Mimicry and the perception of a phonemic contrast induced by silent interval: Electromyographic and acoustic measures. Journal of the Acoustic Society of America, 1961, 33, 842. (Abstract)

Hartvigson, H. A specific case of terminal juncture and syntactic cohesion. Phonetica, 1965, 13, 227-251.

Haugen, E. Phoneme or prosodeme? Language, 1949, 25, 279-282.

Henderson, A., Goldman-Eisler, F., & Skarbek, A. Sequential temporal patterns in spontaneous speech. Language and Speech, 1966, 9(4), 207-216.

Henderson, A., Goldman-Eisler, F., & Skarbek, A. Temporal patterns of cognitive activity and breath control in speech. Language and Speech, 1965, 8, 236-242.

Hill, A. A. Consonant assimilation and juncture in English: A hypothesis. Language, 1955, 31, 533-534.

Hill, A. A. Suprasegmentals, prosodies, prosodemes. Language, 1961, 37, 457-468.

Hill, L. A. Some notes on juncture. Manual of Phonetics, 1956, No. 105, 10-13.

House, A. S. On vowel duration in English. Journal of the Acoustic Society of America, 1961, 33, 9, 1174-1178.

Huggins, A. W. F. Distortion of the temporal pattern of speech by syllable and tied alternation. Language and Speech, 1967, 10, 133-140.

Huggins, A. W. F. The perception of timing in natural speech. Language and Speech, 1968, 11, 1.

Jaffe, J., Cassotta, L., & Feldstein, S. Markovian model of time patterns of speech. Science, 1964, 884, 144.

Johnson, W. Measurements of oral reading and speaking rate and disfluency of adult male and female stutterers and non-stutterers. Journal of Speech and Hearing Disorders, 1961, 1-20(Monogr. Suppl. 7)

Johnson, W., Darley, E. L., & Spriestersbach, D. C. Diagnostic methods in speech pathology. New York: Harper & Row, 1963.

Kasl, S. V., & Mahl, G. F. The relationship of disturbances and hesitations in spontaneous speech to anxiety. Journal of Personality and Social Psychology, 1965, 1, 425-433.

Kelly, J. C., & Steer, M. D. Revised concept of rate. Journal of Speech and Hearing Disorders, 1949, 14, 222-226.

Lallgee, M. G., & Cook, M. An experimental investigation of the function of filled pauses in speech. Language and Speech, 1969, 12, 1, 24-28.

Lane, H. L. Temporal and intensive properties of human vocal responding under a schedule of reinforcement. Journal of the Experimental Analysis of Behavior, 1960, 3, 183-192.

Lane, H. L., & Shrinkman, P. G. Methods and findings in an analysis of a vocal operant. Journal of the Experimental Analysis of Behavior, 1963, 6, 179-188.

Lehiste, I. An acoustic-phonetic study of internal open juncture. Supplement to Phonetica, 1960, 5. Pp. 54.

Lehiste, I. Juncture. In E. Zwirner & E. Bethge (Eds.), Proceedings of the 5th International Congress of Phonetic Sciences. Basil, Switzerland: Karger, 1965. Pp. 172-200.

Liberman, A., Harris, K. S., Eimas, P., Lisker, L., & Bastian, J. An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance. Language and Speech, 1961, 4, 175-195.

Liberman, A. M., Harris, K. S., Kinney, J. A., & Lane, H. The discrimination of relative onset-time of the components of certain speech and non-speech patterns. Journal of Experimental Psychology, 1961, 61, 379.

Lisker, L. Closure duration and the inter-vocalic voiced-voiceless distinction in English. Language, 1957, 33, 42.

Lounsbury, F. G. Pausal, juncture and hesitation. In C. E. Osgood & T. A. Sebeck (Eds.), Psycholinguistics: A survey of theory and research. Baltimore, Md.: Waverley Press, 1954.

Maclay, H., & Osgood, C. E. Hesitation phenomena in spontaneous English speech. Word, 1959, 15, 19.

Mahl, G. Disturbances in the patient's speech as a function of anxiety. Paper read to the Eastern Psychological Association, 1956. (a)

Mahl, G. Disturbances and silences in the patient's speech in psychotherapy. Journal of Abnormal and Social Psychology, 1956, 43, 1-16.

Mahl, G. Normal disturbances in spontaneous speech. Paper read to the American Psychological Association, 1956. (b)

Matarazzo, J. D. Prescribed behavior therapy: Suggestions from clinical research. In A. J. Bachrach (Ed.), Experimental Foundations of Clinical Psychology. New York: 1962, 471.

Minifie, F. An analysis of the durational aspects of connected speech samples by means of an electronic speech duration analyzer. Unpublished doctoral dissertation, State University of Iowa, 1963.

Murray, H. A. Thematic apperception test manual. Harvard College: 1943.

Norwine, A. C., & Murphy, O. J. Characteristic time intervals in telephonic conversation. Bell System Technological Journal, 1938, 17, 281.

O'Malley, M. H., & Peterson, G. E. An experimental method for prosodic analysis. Phonetica, 1966, 15, 1-13.

Peterson, G. E., & Lehiste, I. Duration of syllable nuclei. Journal of the Acoustic Society of America, 1960, 32, 693.

Pickett, J. M., & Pollack, I. Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. Language and Speech, 1963, 6, 151-164.

Pierce, J. E. The supra-segmental phonemes of English. Linguistics, 1966, No. 21, 54-70.

Pittenger, R. E., Hockett, C. F., & Dahehy, J. J. The first 5 minutes: A sample of microscopic interview analysis. Ithaca, N. Y.: P. Martineau, 1960.

Potter, S. Syllabic juncture. In A. Sovijarvi & P. Aalto (Eds.), *Proceedings of the 4th International Congress of Phonetic Sciences*, Helsinki, 1961. The Hague: Mouton & Co., 1962. Pp. 105-119.

Quirk, R. et al. Studies in the correspondence of prosodic to grammatical features in English. In H. Lunt (Ed.), *Proceedings of the 9th Congress of Linguistics*, Cambridge, Mass., 1962. The Hague: Mouton & Co., 1964.

Ramsay, R. W. Speech patterns and personality. *Language and Speech*, 1968, 11, 54-63.

Rischel, J. Stress, juncture and syllabication in phonemic description. In H. Lunt (Ed.), *Proceedings of the 9th Congress of Linguistics*, Cambridge, Mass., 1962. The Hague: Mouton & Co., 1964.

Roberts, P. *Patterns of English*. New York: Harcourt, Brace, 1956.

Roe, A. M., Derbyshire, A. J., & Sedroski, J. Quantifying periods of silence and sound in speech, part 1, non-stuttering compared to stuttering. Paper presented at the convention of the American Speech and Hearing Association, November, 1958.

Schwartz, J., & Jaffe, J. Markovian prediction of sequential temporal patterns in spontaneous speech. *Language and Speech*, 1968, 11, 27-30.

Schwartz, M. F. The lengths of silence in initial s-plosive blends. *Speech Monographs*, 1964, 31, 184-185.

Sharf, D. J. Duration of post-stress intervocalic stops and preceding vowels. *Language and Speech*, 1962, 5, 26-30.

Shapiro, K. *A bibliography of modern prosody*. Baltimore, Md.: Johns Hopkins Press, 1948.

Sharp, A. E. Syllable-division, duremes, and juncture in English. In E. Zwirner & E. Bethge (Eds.), Proceedings of the 5th International Congress of Phonetic Sciences. Basil, Switzerland: Karger, 1965. Pp. 521-524.

Sharp, A. E. The analysis of stress and juncture in English. Transactions of the Philological Society, London, 1960, 104-135.

Sidman, M. Tactics of scientific research; evaluating experimental data in psychology. New York: Basic Books, 1960.

Snidecor, J. C. A comparative study of the pitch and durational characteristics of impromptu speaking and oral reading. Speech Monographs, 1943, 10, 51-56.

Starkweather, J. A. A speech rate meter for vocal behavior analysis. Journal of the Experimental Analysis of Behavior, 1960, 3, 2.

Starkweather, J. A. Vocal behavior: The duration of speech units. Language and Speech, 1959, 2, 146-153.

Stolz, W. A probablistic procedure for grouping words into phrases. Language and Speech, 1965, 8, 219-235.

Swadesh, M. On the analysis of English syllabics. Language, 1947, 23, 137-150.

Tannenbaum, P., Williams, F., & Hillier, C. Word predictability in the environment of hesitation. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 134.

Tannenbaum, P. H., Williams, F., & Wood, B. S. Hesitation phenomena and related encoding characteristics in speech and typewriting. Language and Speech, 1967, 10, 203-205.

Trager, G. L. Some thoughts on "juncture". Studies in Linguistics, 1963, 16, 11-22.

Twaddell, W. F. Stetson's model and the "supra-segmental phonemes". Language, 1953, 29, 415-53

Verzeano, M. Time patterns of speech in normal subjects: I. Journal of Speech and Hearing Disorders, 1950, 15, 197-201.

Verzeano, M. Time patterns of speech in normal subjects: II. Journal of Speech and Hearing Disorders, 1951, 16, 345-350.

Verzeano, M., & Finesinger, J. E. An automatic analyzer for the study of speech in interaction and in free association. Science, 1949, 110, 45.

Webb, J. T. Subject speech rates as a function of interviewer behavior. Language and Speech, 1969, 12, 1.

Weinrich, H. V. Phonologic der Sprechpause. (Phonology of speech pauses.) Phonetica, 1961, 7, 4-18.

DURATIONAL CHARACTERISTICS OF SPONTANEOUS SPEECH:
EFFECTS OF SAMPLE SIZE ON TEST-RETEST RELIABILITY

by

JANICE USTRUD THOMS

B. A., Augustana College, 1964

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

The purpose of this research was to demonstrate test-retest reliability of durational characteristics of speech and silence events as a function of sample length during spontaneous speaking. Ten young adult males showing no speech or hearing pathology were the subjects for the study. The spontaneous speech of these subjects was elicited by the picture Thematic Apperception Test, and recorded on two occasions separated by 24 hours.

Speech was described as periods of silence events interspersed with periods of speech events. Frequency distributions for these speech and silence events were derived from recorded speech displayed on a cathode ray oscilloscope. The speech and silence events were photographed, measured and tabulated by duration into appropriate class intervals. The number of speech events within each interval was subtotaled within time units of 30 seconds and cumulated at 5 minutes for trial I and then for trial II. Percentage of individual test-retest reliability was measured at each class interval for each 30 second time unit. Median values were then computed by ranking individual percentages of reliability at each class interval for one minute units and performing the appropriate computations.

Reliability was seen to peak at 3 minutes and thereafter to remain relatively stationary for four of the eight speech intervals and for three of the silence intervals. Average reliability peaked at a mean time of 3.5 minutes for speech intervals and at a mean time of 3.9 minutes for silence intervals. The tendency was observed for reliability of durations of speech and silence events to increase with sample length increase and then to stabilize after the point of highest reliability was attained. Previous research findings indicating that durations of silence events are more variable than speech events were confirmed in this study.

The observed steady states of the durational characteristics under present study indicate that a 4 minute sample of speech represents an adequate gathering of spontaneous speech behavior from which to establish frequency distributions for duration of speech and silence events.