An Investigation of Isochronism in the Rhythm of American English Speech.

Domitila Domenech De belaval

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/1143
AN INVESTIGATION OF ISOCRONISM IN THE RHYTHM OF

AMERICAN ENGLISH SPEECH

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Program in Linguistics

by

Domitila Domenech de Belaval
B.S., Elmira College, 1933
M.A., Columbia University, 1947
June, 1966
ACKNOWLEDGMENT

Acknowledgment of sincere appreciation is due to Dr. George H. Gunn and to Dr. Claude L. Shaver for their encouragement, guidance, and help in the preparation of this study.
TABLE OF CONTENTS

CHAPTER I. INTRODUCTION 1
   The problem 2
   Definition of terms 2
   Review of the literature 8
   Restatement of the problem 24

CHAPTER II. PROCEDURE 25
   Validation of the investigator's judgment of stress 26
   Procedure followed in obtaining corpus 29
   Procedure followed in obtaining data on impromptu speech 30
   Spectrographic analysis of the data 31
   Procedure followed in obtaining data on structured speech 41

CHAPTER III. RESULTS 50
   Interpretation of data on impromptu speech 50
   Interpretation of data on structured speech 53

CHAPTER IV. CONCLUSIONS 63
   Summary 63
   General conclusions 66
   Suggestions for further studies 68

BIBLIOGRAPHY 70
APPENDIX 82
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Percentage of Agreement between Listeners and Consensus</td>
<td>28</td>
</tr>
<tr>
<td>II</td>
<td>Preliminary Values of Linear and Curvilinear Effects in the Relation between Number of Syllables and Duration</td>
<td>51</td>
</tr>
<tr>
<td>III</td>
<td>Significance of Linear and Curvilinear Effects in the Relation between Number of Syllables and Duration</td>
<td>52</td>
</tr>
<tr>
<td>IV</td>
<td>Coefficients of Correlation between Number of Syllables and Duration</td>
<td>53</td>
</tr>
<tr>
<td>V</td>
<td>Observed Duration of Framed Utterances</td>
<td>54</td>
</tr>
<tr>
<td>VI</td>
<td>Comparison of the Number of Accurate Predictions of Duration in Terms of the Theory of Isochronism and in Terms of a Hypothesis of Linearity</td>
<td>56</td>
</tr>
<tr>
<td>VII</td>
<td>Comparison of the Mean Average Percentage of Prediction Accuracy in Terms of the Theory of Isochronism and in Terms of a Hypothesis of Linearity</td>
<td>57</td>
</tr>
<tr>
<td>VIII</td>
<td>Average Duration of Syllables in 180 Samples of Stress Groups from the Free Speech of All Subjects</td>
<td>59</td>
</tr>
<tr>
<td>IX</td>
<td>Percentage of Equal or Approximately Equal Adjacent Stress Groups in Samples from the Free Speech of All Subjects</td>
<td>60</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGURE 1</td>
<td>Segmentation of Adjacent Two-Syllable (30 cs) and Six-Syllable (87 cs) Stretches Spoken by Subject 4</td>
<td>35</td>
</tr>
<tr>
<td>FIGURE 2</td>
<td>Segmentation of Three-Syllable Stretch Measuring 51 cs</td>
<td>36</td>
</tr>
<tr>
<td>FIGURE 3</td>
<td>A Seven-Syllable Stretch (120 cs) Adjacent to a Three-Syllable Stretch Spoken by Subject 3</td>
<td>37</td>
</tr>
<tr>
<td>FIGURE 4</td>
<td>Broad Band Spectrogram of a One-Syllable Stretch. Subject 3</td>
<td>38</td>
</tr>
<tr>
<td>FIGURE 5</td>
<td>Narrow Band Spectrogram. One-Syllable Stretch Shown in Figure 4. Duration—22 cs</td>
<td>39</td>
</tr>
<tr>
<td>FIGURE 6</td>
<td>Duration of 33 Samples of Stress Groups in Impromptu Speech. Subject 1</td>
<td>43</td>
</tr>
<tr>
<td>FIGURE 7</td>
<td>Duration of 31 Samples of Stress Groups in Impromptu Speech. Subject 2</td>
<td>44</td>
</tr>
<tr>
<td>FIGURE 8</td>
<td>Duration of 30 Samples of Stress Groups in Impromptu Speech. Subject 3</td>
<td>45</td>
</tr>
<tr>
<td>FIGURE 9</td>
<td>Duration of 32 Samples of Stress Groups in Impromptu Speech. Subject 4</td>
<td>46</td>
</tr>
<tr>
<td>FIGURE 10</td>
<td>Duration of 33 Samples of Stress Groups in Impromptu Speech. Subject 5</td>
<td>47</td>
</tr>
<tr>
<td>FIGURE 11</td>
<td>Duration of 35 Samples of Stress Groups in Impromptu Speech. Subject 6</td>
<td>48</td>
</tr>
<tr>
<td>FIGURE 12</td>
<td>Combined Results. 180 Samples</td>
<td>49</td>
</tr>
<tr>
<td>FIGURE 13</td>
<td>Duration of Adjacent Stress Groups in Free Speech. Subject 1</td>
<td>61</td>
</tr>
<tr>
<td>FIGURE 14</td>
<td>Duration of Adjacent Stress Groups in Free Speech. Subject 4</td>
<td>62</td>
</tr>
</tbody>
</table>
ABSTRACT

This study attempts to test the validity of the alleged periodicity of stress groups in the rhythm of American English speech. The results of this study contradict the conclusion reached by Classe in 1939, that English prose rhythm is essentially isochronic. They confirm in part the observation made in 1962 by Shen and Peterson, who found no factual basis for the theory of isochronic rhythm. In addition, the results serve as the basis for a tentative hypothesis of linear relation between number of syllables and time.

The initial corpus for this study consists of 1200 samples of stress groups selected from 10-minute magnetic tape recordings of impromptu speech by each of six young males, all speakers of North Midland dialect of Standard American English. It also includes 128 recordings of examples of isochronic rhythm extracted from textbooks. Two hundred examples of uninterrupted stretches of free speech containing from one to nine syllables and delimited by the onset of strong stress were chosen from each recording. From these, four or, in some cases, fewer samples of each type were randomly selected and spectrographed. Syllable centers and major stresses were identified on each spectrogram. The syllables in each stretch were counted and the duration of the stretch was measured from the onset of one stressed vowel to the onset of the next stressed vowel. Frequency distributions were plotted for each subject and for the combined subjects. Since the frequency distributions showed a possible linear or curvilinear relation between the number of syllables and the duration of the stretches, the data were submitted to mathematical analysis in order to determine
the significance of these effects. Application of equations for curvilinear effect and for linear regression, and a t-test show a definite linear relation, in every case with a correlation coefficient above 0.90, for each subject and for the combined subjects. On the basis of these findings, a tentative hypothesis of direct linear relationship between number of syllables and duration is proposed.

Sets of framed utterances similarly rendered by three or more subjects were measured from absolute beginning to absolute end, and the duration of the utterances was compared. In 95% of the comparisons the duration of an utterance containing more syllables was found to be greater than the duration of a paired utterance containing fewer syllables. Comparison of the accuracy of predicting the duration of these paired utterances on a basis of theoretical isochronic rhythm and on a basis of hypothetical linearity shows that those comparisons based on a hypothetical linearity are decidedly more accurate both in terms of number of predictions and in terms of percentage of predictive accuracy.

In view of these results the conclusion is reached that the theory which assumes that, because of the essentially isochronic character of English rhythm, the duration of stress groups is independent of the number of unstressed syllables contained between strong stresses is not supported by factual evidence.

A cursory look at the unequal durations of abutting stress groups found incidentally in the samples of free speech constituting the corpus of this investigation points to the need for further study of this aspect in order to confirm or disprove the conclusions here reached.
CHAPTER I. INTRODUCTION

The rhythm of English speech is usually described as isochronic; that is, a rhythm characterized by the recurrence of stress at more or less uniform intervals of time without regard to the number of syllables between stresses. Many textbooks, especially those devoted to the teaching of English as a foreign language, state this general belief as a rule for the correct pronunciation of the language. Examples, such as Pike's traditional one,

\[
\begin{align*}
\text{English is easy.} \\
\text{English is very easy.} \\
\text{The English lesson is very easy.}
\end{align*}
\]

illustrate the rule. In this set of sentences the lapse of time between the stressed syllables of English and of easy is supposed to be the same.

Teachers of English as a second language have at times doubted the validity of this assertion. They find that prose passages read with isochronic rhythm often sound mechanical, and that exercises meant to be read with isochronic rhythm must be read with an unnatural tempo. An example of these exercises is Prator's

\[
\begin{align*}
\text{Boys need money.} \\
\text{The boys will be needing some of their money.}
\end{align*}
\]

In addition, the construction of drills that would accurately illustrate the principle of isochronic rhythm is extremely laborious. If true examples are rare, can isochronism be an essential characteristic of the rhythm of spoken English?
The Problem

The purpose of this study is to investigate the assumption that the rhythm of spoken English is essentially isochronic; that is, to ascertain whether in impromptu speech the duration of intervals between stresses is independent of the number of syllables between those stresses. Before an investigation of this problem can be conducted or any conclusions reached, such debatable terms as rhythm, stress, and syllable must be clarified, and the meaning of each limited to that used in this study.

Definition of Terms

Rhythm. To Aristotle (2) rhythm was an essential element of speech, its "measured movement and proportion" attracting attention to the orator. Many of the current definitions of speech rhythm may be viewed as attempts to specify the characteristics of this "measured movement and proportion." According to Webster's New and Third New International Dictionary of the English Language (164, 165), the rhythm of speech consists of the regular recurrence of similar features, or of the ordered alternation of strong and weak elements. In the Oxford English Dictionary (114) the rhythm of speech is defined as the "measured recurrence of certain rhetorical features," determined by quantity (meaning duration) or stress, or both. In this definition it may be noted that there is an uncertainty as to whether recognition should be given predominantly to the succession of time-spans which recur regularly or to the repetition of features which establish the boundaries of time-spans.
A student of linguistics who has some knowledge of acoustics and
of psycho-acoustics might be inclined to define the rhythm of speech as
the patterned repetition of certain acoustic phenomena. To this per-
son, the acoustic phenomena might be differentiable changes in inten-
sity, pitch, duration, or quality, alone or in any combination. To this
person, the question of whether a perceived rhythm is related to the
recurrence of the acoustic phenomena which demarcate the time limits
of a sound pattern or to the duration of silence or relative silence
within given boundaries is subject to choice.

In this study the acoustic phenomena chosen as determinators of
speech rhythm are the lapses of time between successive boundaries
which have been judged as stress points. The assumption that rhythm
is the result of the recurrence of intervals of speech sounds between
successive stress points is, then, the basis of this study.

Stress. Since there is, as yet, no conclusive evidence that
stress is concomitant with any one acoustic phenomenon, stress is
regarded, for the purpose of this investigation, as the emphasis placed
on the nucleus of a syllable by means of intensity, duration, or change
in fundamental frequency, any or all of which, when perceived by the
hearer, may be interpreted as stress\(^{(8, 17, 49, 50, 53, 129)}\).

Not all linguists, however, agree that these three factors are the
most essential in stress. Some correlate stress primarily with inten-
sity\(^{(115, 157)}\). Others correlate it with fundamental frequency\(^{(23, 110, 112)}\). Some combine intensity and fundamental frequency\(^{(73, 131)}\).
For Tiffin and Steer\(^{(156)}\), Carrell and Tiffany\(^{(20)}\), and Lieberman\(^{(97)}\),
duration is the most critical factor, whereas for Long\(^{(99)}\), both
duration and fundamental frequency are critical. Fletcher\textsuperscript{(45)} claims that formant (overtone) structure is as essential as intensity or fundamental frequency. These three factors, as well as duration, are mentioned by both Firth\textsuperscript{(41)} and Heffner\textsuperscript{(66)} as possible features of stress. Frequency modulation is proposed by Bolinger\textsuperscript{(15)} as the decisive element; and Voelker\textsuperscript{(161)}, Ortleb\textsuperscript{(114)}, and Jones\textsuperscript{(79)} add it to the other properties, making a total of five: intensity, duration, fundamental frequency, formant structure, and frequency modulation. On the other hand, some linguists suggest that stress should not be analyzed in terms of acoustic properties but primarily in terms of muscular effort\textsuperscript{(17, 79)}. Lehiste and Peterson\textsuperscript{(94)}, and Ladefoged, Draper, and Whitteridge\textsuperscript{(89)} claim it is perceived primarily by kinesthetic memory. And, finally, there are those who believe that stress, being essentially subjective, cannot be measured by instruments at all\textsuperscript{(22, 106)}.

All linguists agree that degrees of stress are relative, but many disagree on the number of degrees it may have. Some claim that there are an infinite number of degrees, or, at any rate, as many as there are syllables in a word\textsuperscript{(71, 88, 150)}. Others, notably Martinet\textsuperscript{(105)} and Arnold\textsuperscript{(4)}, assert that there are only two essential degrees, stress and unstress, any intermediate stresses being either reduced variants of strong stress or "non- tonic" strong stress. A 6-step scale of stress gradation was proposed by Liddell and accepted by Morris\textsuperscript{(109)} in 1936. Jones\textsuperscript{(81)} and Lado\textsuperscript{(92)} have recently proposed 5-step scales. Most American phonemicists, among them Block\textsuperscript{(11)}, Trager\textsuperscript{(12)}, Smith\textsuperscript{(158)}, Hill\textsuperscript{(67)}, Gleason\textsuperscript{(52)}, and Hockett\textsuperscript{(70)}, prefer a 4-step scale, while phoneticians generally
prefer a 3-step scale(3, 17, 73, 82, 86, 119, 160).

Since isochronic rhythm is usually described as a succession of
similar intervals between primary or dominant stresses, a distinction
of only two degrees of stress (strong and weak) is sufficient for the pur-
pose of this study. Strong stress will include the most prominent
 stresses in an uninterrupted stretch of speech, that is, those stresses
that make a syllable stand out above all other syllables in the stretch.
These are the phonetic primary and level stresses, and the phonemic
primary stress. Weak stress will include not only inconspicuous
 stresses but also phonetic secondary and phonemic secondary and ter-
tiary. This two-fold division of stress is not unusual in rapid speech,
during which medium stresses are considerably weakened, and poten-
tial stresses are subordinated and unrecognized for rhythmic purposes
(4, 92, 105, 120). In fact, the binary stress scale may soon be the sub-
ject of much speculation since, according to Garvin and Trager\(^{51}\),
equipment for machine translation into and from English is expected in
the near future to be programmed for the recognition of only two
degrees of stress.

The syllable. Though the oldest, commonest, and often easiest
analysis of the stream of speech seems to be by division into syllables,
the nature itself of the syllable and the identification of its boundaries
have long been baffling linguistic problems. Some scholars, notably
Z. Harris\(^{63}\), Block and Trager\(^{12}\), Trager and Smith\(^{158}\), and
Von Essen\(^{162}\), have denied the physical reality of the syllable.
Others have considered it a separate perceptual phenomenon which
may or may not have a basis in physical reality\(^{10, 11, 25, 70}\).
Still others have suggested that it may consist mostly of kinesthetic sensations, stimulated by alternations of tension and relaxation \(^{(27, 28, 53, 127)}\). Spectrographic evidence obtained by Lehiste \(^{(93)}\) seems to point to the syllable as a phonetic sequence bounded by certain marginal variants of sounds; on the other hand, Hjelmslev \(^{(69)}\) and Whitehall \(^{(167)}\) have described it as merely a carrier of prosodic features. Most scholars, however, agree that the syllable is an undeniable unit of the language although its nature may vary according to the analytical viewpoint from which it is described.

On the basis of acoustic criteria, the prevailing opinion is that the syllable is a speech unit containing a peak or crest made prominent by greater sonority (sound pressure, intensity, or output of energy). At times, however, increased fundamental frequency, longer duration, or a combination of these may determine the peak. Less sonorous margins or slopes, contrasting with the peak, form the remainder of the syllabic unit. Since the most sonorous sounds are vowels, liquids, and nasals, in that order, one of these sounds ordinarily constitutes the peak \(^{(10, 13, 66, 81, 86, 88, 121, 150, 160)}\). A point of minimum sonority (or dip in output of energy) determines the boundary between successive syllables, though this point may not always be determinable \(^{(20, 71, 77, 78, 81, 86, 121, 160)}\). For the purpose of this study, the nature of which requires that syllables be counted, the following assumptions are made on the basis of the preceding paragraphs:

1. that syllables are perceptible, and therefore countable, successive peaks of prominence in the stream of speech; and,
2. that syllable centers are acoustically identifiable, and therefore countable, vowel or vowel-like segments contrasting, by reason of greater intensity or longer duration, with adjacent consonantal segments.

These assumptions do not by any means preclude other generally accepted concepts of the syllable as a physiological or structural unit.

Viewed as an articulatory phenomenon, the concept of the syllable proposed by Stetson in 1928 has prevailed until recently. Stetson described the syllable as a correlate of a single chest pulse, that is, a unit of sounds produced by the volume of air forced upward through the vocal channel by a single ballistic (sharp, non-gradual) compression of the intercostal muscles. The division between syllables, then, would be brought about by an arresting movement of these muscles. As a result, syllables are as discrete, additive, and countable as chest pulses. Among the many phoneticians who have agreed with Stetson's concept of the syllable are Gray and Wise, Davis, Heffner, and Black. Pike and Gleason have agreed in part, the latter accepting Stetson's theory with certain reservations and as applicable only to the phonetic syllable. In recent years, however, other phoneticians have questioned the chest-pulse theory. In 1958 Ladefoged, Draper, and Whitteridge claimed that experiments performed by them often showed no correlation between bursts of muscular activity and the number of segments perceived as syllables. In the same year, Skalickova described the syllable as the simplest and closest articulatory, but not pulmonary, unit. And in 1964, Hála proposed that the indispensable physiological basis of the syllable center, and
hence of the syllable itself, is vocal cord vibration.

McQuown's(105) suggestion, that greater sonority may be the result rather than the cause of a sound's being used as a syllabic, reflects the attitude of many linguists who do not consider the syllable necessarily as a phonetic unit, but primarily as a form unit in a system. According to Hjelmslev(69), Block(11), and Hockett(70), the syllable is a structural unit which plays a role in the formation of larger units and is determinable only by phonemic analysis. Haugen (65) says it is non-contrastive and does not bear meaning. Whitehall (167), Pike(120), Bendor(7), and others(64, 142, 151) describe it as a recurrent sequence of segmental phonemes containing a nucleus which serves as a unit of placement for stress as well as for other prosodic features, such as tone, length, rhythmic grouping, juncture markers, or morpheme structure. The nucleus, which is vocalic and irreducible, stands out against consonantal margins(52, 67, 76, 99). Syllables are discrete and countable entities even though their boundaries may overlap and be undefinable, except on the higher levels of analysis (42, 43, 84). As Haugen(64, 65) and Black(9) have pointed out, the fact that pauses, hesitations, and interruptions in the stream of speech occur, not within syllables but between them, marks syllables as undeniable units of the language.

Review of the Literature

In 1816, when Coleridge published "Christabel" together with an explanation of the accentual character of English rhythm, he started a controversy which led indirectly to the subject of this investigation.
The numerous impressionistic studies of English verse rhythm that followed his in turn aroused interest in the rhythm of prose and of everyday speech. Experimental work, however, on the rhythm of English began in the last quarter of the 19th century. From its findings two principal theories gradually developed. Their proponents may be called the "timers" and the "non-timers."

The "non-timers," manifestly aware of irregularities in number of syllables and consequently in time-lapses between rhythmic beats, considered duration a secondary element in the rhythm of speech. They proposed other factors, most frequently accent, as its essential element. In 1882 Guest\(^{(56)}\) claimed that accent, independent of quantity, is the sole principle of English rhythm. Around the turn of the century several other interpretations were proposed. Liddell\(^{(97)}\) postulated "waves of impulse," based on accent, as the fundamental units of rhythm, while Hurst and McKay\(^{(74)}\) contended that both time and accent might occasionally be ignored, allowing thought alone to regulate rhythm. According to Ebbinghaus\(^{(36)}\), rhythm is a succession of unified groups of sensations and not the mere succession of impressions following one another at equal intervals of time. MacDougall\(^{(102)}\) insisted that accentuation as well as recurrence and rate is essential to a rhythmic impression. He added that, although the sense of temporal equivalence is a necessary element, there is never any repetition of identical sequences.

In 1901 Wallin\(^{(163)}\) published the results of his research on the rhythm of speech. As a marker of the rhythmic beat he proposed the centroid, an objective physical phenomenon measurable in amplitude,
period, and length, and simultaneously an event in consciousness; in
other words, an emphatic syllable. Careful measurements of centroid-
intervals (intervals containing one strong syllable, any number of inter-
vening syllables or none, and no pauses) led him to the conclusion that
the length of the average centroid-interval is invariably proportional to
the number of syllables that compose it. Though agreeing with Wallin
on the role of the centroid as constituting the rhythmic beat and on the
importance of pitch as well as accent in the determination of centroids,
Scripture\(^{(136, 137)}\) emphasized that the subjective impression of
equal time-intervals is more important than their objective values.
Lipsky\(^{(98)}\) agreed that rhythmical units certainly do not follow each
other with regularity; but, he argued, since equal time-intervals are
sometimes filled with unequal numbers of unaccented syllables spoken
more rapidly than accented syllables, irregularities in time-intervals
may occur without destroying the rhythm. Thus, only approximate
equality of time-lapses is necessary in prose rhythm.

F. N. Scott\(^{(132)}\) claimed that the fundamental element in speech
appeared to be the inflectional arc, an upward glide followed by a
downward glide in pitch. Fijn van Draat\(^{(40)}\) asserted that prose
rhythm is a movement measured by the regular alternation of accented
and unaccented syllables, but Elton\(^{(37)}\) found their recurrence and
alternation most irregular and insisted that he could not hear equal
measures. In \textit{A History of English Prose Rhythm}, Saintsbury\(^{(126)}\)
declared himself against any systematic theory and refused to discuss
the nature of rhythm at all because of the numerous variations in its
time-units.
The "timers" maintained that the element of time is most important in rhythm, that emphasis is subordinate, and that the sounds of a rhythmical sequence must be periodic. In 1894 Bolton\(^{(16)}\) noted the tendency to hasten or to slow up the elements in a group in order to make the group fit a natural periodicity. In his earlier experiments Sievers\(^{(139)}\) found prose divided into sections of approximately equal duration. Later he included "stress-gradation" together with "time-organization" as the most important factors in rhythm\(^{(140)}\). Both Meumann\(^{(107)}\) and Miner\(^{(108)}\) agreed that rhythmic units of prose, though not following each other with perfect regularity, tend to be similar in length; and that equal intervals of time may be filled with unequal numbers of unaccented syllables, the larger number being spoken more rapidly. Alden\(^{(1)}\) suggested that the fundamental principle of English prose rhythm, in which accents appear at regular intervals, is the basis of English verse rhythm.

In 1909, in an essay on English metrics, Verrier\(^{(159)}\) used the term isochronic to describe French rhythm as contrasted with English rhythm. He found, nevertheless, in both English verse and prose, an unconscious tendency to bring the consecutive rhythmic segments to an equal duration, even though their irregularities might lead to the conclusion that equality of time-intervals is an illusion. Shortly afterwards, Warner Brown\(^{(18)}\), too, stressed the subjective impression of temporal regularity as the only undisputed characteristic of rhythm; and MacColl\(^{(101)}\) went so far as to claim that accents really remain equidistant in time, though extra-rhythmical variations may occur, thereby enriching the rhythm and relieving the monotony of prose.
Fourteen years after Verrier's important study of English metrics, Thomson\(^{(149)}\) asserted in his extensive book, *The Rhythm of Speech*, that the basis of English prose rhythm is timing, that every syllable has a numerical value, and that the sums of all individual syllables between accents (or other marking points) are equal. Thus, by 1923, the theory of isochronic prose rhythm was well defined though as yet not identified by name. * Two years later J. H. Scott\(^{(133)}\) elaborated on Thomson's theory by adding two important observations; namely, that to strike an average in the timing of phrasal parts, the tempo of speech increases in direct proportion to the number of syllables in each phrase; and that the phrasal parts may be "equal, or approximately equal, or apparently approximately equal." And in 1926 Armstrong and Ward\(^{(3)}\) described and illustrated isochronic rhythm in the following way:

> In each group the stressed syllables occur at more or less regular intervals of time; and the unstressed syllables, whether many or few, occupy the time between the stresses; it is this factor that gives English its characteristic rhythm. 

> These three sentences,

> I don't think I can do it.
> I shouldn't of thought he could do it.
> I shouldn't of thought it possible to do it.

> each with the same number of stressed syllables, but with a different number of unstressed syllables, may be said in the same time, although the number of syllables between the first two stresses is greater in each sentence than in the one that precedes it.

> In *The Rhythm of Prose*, published in 1916, Patterson\(^{(116)}\) tried to bridge the gap between the "timers" and the "non-timers" by

\*The first use of the term *isochronic* to describe English prose rhythm seems to have been made by André Classe in 1939. Cf. pages 13 ff.
substituting for the old concept of objective equality of rhythmic units, one of subjective equality. This was not a new concept by any means. Miner (108), Woodrow (169), and Wundt (170) had previously pointed out the illusory character of time-images, which did not correlate directly with sensations from objective stimuli but were overestimated or underestimated under the influence of "emotional tone." This subjective quality of rhythm was again stressed by Lotspeich (100) in 1922 when, in an attempt to differentiate between the rhythm of poetry and that of prose, he noted that in prose the regular recurrence of stress was interrupted by its intellectual content.

Only two thorough attempts to investigate objectively the theory that English prose rhythm is isochronic seem to have been made before this study. The results of the first investigation, published by André Classe in 1939, favor the theory; the results of the second, made by Yao Shen and Giles Peterson in 1962, refute it.

Classe (22) set out to measure physical time on kymographic tracings of framed utterances read by subjects of varied sex, age, place of birth, education, and profession. All except one were speakers of "approximately British Standard Received dialect." The subjects read prosaic, unemotional, especially structured sentences with which they were well-acquainted. To enhance the rhythm, the subjects tapped out the rhythmic beats on a telegraph key as they read. Considering each sentence as a rhythmic unit, Classe excluded from the analysis all initial and final unaccented syllables as unessential to the rhythmic scheme. Then he measured adjacent stretches between successive stress points. For practical reasons (ease of segmentation on
kymographic tracings), Classe selected as a fixed point of measure-
ment the moment of explosion of voiceless plosives before stressed
vowels, and, after other consonants, the moment just before the begin-
ning of the stressed vowel. Examples of the two adjacent stress groups
measured in some sentences are given below.

We had / biscuits, / muffins, / tea, and coffee.
He / ran, / jumped over it, / fell, and shouted.
He was a / man of / manifold ac / tivities.
We'll / start direc / tly you're / ready. [sic]

A comparison of the mean durations and standard deviations of the
stress groups in these readings showed that many of the adjacent
groups were not equal. As a result, Classe concluded:

1. that although isochronism is probably the essential
   characteristic of English prose rhythm, it is equally
certain that perfect isochronism can only be realized
when definite conditions are fulfilled;
2. that the conditions necessary for perfect isochronism
   are "similarity of phonetic structure, similarity of
   grammatical structure of the groups, and similarity
   of connection between the groups";
3. that these conditions are comparatively seldom met
   with in ordinary speech and in ordinary kinds of prose;
4. that only groups of 2, 3, and, at most, 4 syllables form
   easily recognizable rhythmic patterns.

Classe also analyzed readings of five passages of different
types of literary prose (from the Song of Songs, Ruskin, Thackeray,
Shaw, and Charles Morgan), and compared the duration of successive
stress-groups and the number of syllables in each. Although he found
the experimental method impractical for the analysis of long passages,
he reached the following conclusions:

1. that in most cases there is a significant relation
   between the number of syllables and the duration of
   stress-groups, though the tendency to level the quantity
   of component syllables is always present;
2. that even in careful prose the accents seldom seem to
   return at rigorously isochronous intervals.
Despite these observations, Classe asserted that an English sentence is normally composed of more or less isochronous groups which contain a varying number of syllables; that stress-groups containing different numbers of syllables tend, everything else being equal, to assume approximately the same duration; that the increased acceleration of longer groups is a result of the tendency toward perfect isochronism; and that speakers probably tend to place stresses in such a way as to facilitate the perception of groups as equal in duration.

In The Intonation of American English, Pike\(^{(119)}\) acknowledged the influence of Classe's observations and wrote:

The timing of rhythm units produces a rhythmic succession which is an extremely important characteristic of English phonological structure. The units tend to follow one another in such a way that the lapse of time between the beginning of their prominent syllables is somewhat uniform. Notice the more or less equal lapses of time between stresses in the sentence The 'teacher is 'interested in 'buying some 'books; compare the timing with the following and notice the similarity in that respect despite the different number of syllables: 'Big 'battles are 'fought 'daily . . . The tendency toward uniform spacing of stresses in material which has an uneven number of syllables within its rhythm groups can be achieved only by destroying any possibility of even time-spacing of syllables. Since rhythm units have different numbers of syllables, but a similar time value, the syllables of the longer ones are crushed together, and pronounced very rapidly, in order to get them pronounced at all, within that time limitation . . . . Compare the similar timing and stresses but variant number of syllables in the following pairs of illustrations:

The 'man's / 'here. If 'Tom will / 'I will.

The 'manager's / 'here. If 'Tom'll do it'/ 'I will.

Two years later, in Phonemics, Pike\(^{(120)}\) added:

In English one tends to hear stress-timed rhythm in contra-distinction to a syllable-timed rhythm . . . . Since in this rhythmic style there is a tendency for the stressed syllables to occur at more or less evenly spaced intervals in time, and since between two such stresses there may be an
indeterminate number of syllables, then if only one syllable comes between the stresses, it will be likely to receive more time, emphasis, and clarity than will three or four unstressed syllables if they all occur between the same two stresses. Compare these examples: The 'teacher 'came. The 'teacher is the one who 'came. If the length of time between the stressed syllables teach- and came is kept the same, then the syllables -er is the one who are more rapid than the -er of the first sentence.

After Pike's apparently definitive statements, almost all authors touching on the timing of English prose rhythm have reflected his opinion. Thus, in an inquiry into English prosody, Seymour Chatman(21) remarked that a very important feature of English is its isochronism, that is, its tendency to squeeze units into relatively equal time-spans marked by stress pulses.

In 1956 Whitehall(166), in a review of Trager and Smith's Outline of English Structure, called attention to a certain feature of English,

one not mentioned in the Outline and one not directly a significant part of English structure. This is the fact, first noticed by Pike, that the time-lapse between any two primary stresses tends to be the same irrespective of the number of syllables and of the junctures between them. In short, unlike such 'syllable-timed' languages as Spanish, English is 'stress-timed' or isochronic. . . . Isochronism is produced not only by accelerating and crushing together the syllables between primary stresses but also by increasing or decreasing the pauses which always may follow the three terminal junctures.

Two years later Whitehall and Hill (168) reported the following:

This feature of English is the fact that the amount of time between two primary stresses tends to be the same, irrespective of the amount of material between them. This feature can be called isochronism, and it is intimately related to juncture, since isochronism is often secured by increasing or decreasing the pause which always may accompany a terminal juncture. In contrast to English, such a language as Spanish tends to make all syllables occupy the same amount of time, thus producing the effect we have called isosyllabism. . . . The Pike drill sentences,
English / is easy
English / is very easy
The English lesson / is very easy

will demonstrate what is meant.

Robert Lado\(^{(92)}\) described English prose rhythm as follows:

English stress rhythm is characterized by a primary stress in each phrase, and accompanying secondary and weak stresses, with a tendency to achieve approximately the same length of time for each phrase regardless of the number of syllables involved. ... This emphasis [sentence stress] and the length of stressed syllables and the tendency to uniform length between stresses makes English rhythm "phrase timed" rather than "syllable timed."

Comparing the sound-systems of English and Spanish, he wrote:

English rhythm, in contrast, is phrase-centered. In English the phrases tend to be of the same length, and the syllables are shortened or stretched to fit the uniform length of the phrases\(^{(90)}\).

In Language Testing he put it this way:

Each language has certain permitted sequences of stresses and length units which, as they follow one another, give it a rhythm of its own. English tends to make uniform the length of time from primary stress to primary stress regardless of the number of intervening syllables. Thus the phrases separated by a vertical line in the following example are of approximately the same duration even though the first has only one syllable and the second has four.

Sign | the documents | in ink\(^{(91)}\).

Discussing the founding of English metre, John Thompson\(^{(152)}\) claimed that the rhythm of verse imitates the rhythm of speech, in which the stresses tend to occur at isochronous intervals. And James W. Harris\(^{(62)}\), in an unpublished master's thesis presented at Louisiana State University in 1962, asserted that in English the stronger stresses in an utterance are usually separated by about equal lengths of time, no matter how many syllables with weaker stresses may intervene.
In a brief discussion of rhythm in The Phoneme, Daniel Jones (80) expressed this opinion:

In stress languages [among which he had previously placed English], there is usually a tendency to make strong stresses follow each other at fairly equal intervals, whenever this can conveniently be done. This tendency produces the effect commonly called "rhythm." . . . If a number of weakly stressed syllables intervene between a strong stress and the next following strong stress in a sentence, various shortenings may take place in an unconscious endeavour to make the "stress bar" equal in length to other "stress bars."

Not all textbooks dealing with phonology discuss rhythm. Most of those that do, however, describe English rhythm as isochronic. In A Manual of Phonology(71), Hockett wrote: "In English the higher stresses in an utterance are normally separated by about equal lengths of time, no matter how many syllables with lower stresses may intervene." He pointed out that in these two sentences,

Find a board for me.
Interpret this poetry for me.

the interval between find and board is not much shorter than that between -ter and po-. Again, in A Course in Modern Linguistics, Hockett (70) made the same point twice. Describing English rhythm as stress-timed, he said:

This means that it takes about the same length of time to get from one primary-stressed syllable to the next, in speaking at a given overall tempo, whether there are no syllables between them or many. If there are none, we slow down our rate of speech slightly; if there are many we squeeze them in fast . . . . The typical timing of two examples can be shown as follows:

The | wind | blèw up the | strèet.
The | wind blèw | up | the | strèet.

Later in the book, he expressed it this way:

English has stress-timed rhythm: the length of time it takes to produce an utterance depends roughly on the number of
stresses in it, so that unstressed syllables are sometimes squeezed together and produced very rapidly, if there are a number of them between two successive stressed syllables.

Bronstein\(^{(17)}\) put it briefly: "The rhythm pattern of our speech may be noticed in the tendency for stresses to recur at regular intervals in sentences as well as in words."

Prator\(^{(126)}\) discussed isochronic rhythm extensively. He wrote:

Accents tend to recur at regular intervals. The more unstressed syllables there are between accents, the more rapidly (and indistinctly) they are pronounced. This is true to a large extent even of prose.

In the directions for a set of exercises, Prator pointed out that the addition of extra syllables in successive sentences would not appreciably lengthen their duration because they all had the same number of stresses. He then asked the student to tap out the beat in the first sentence of each set, to pronounce the successive sentences according to the beat, and to bring in all unstressed syllables between the beats.

Some of the sentences from two of the sets are given below.

Dógs éat bónes.
The dógs will éat the bónes.
The dógs will have éaten the bónes.

The cár is héré nów.
The cár will be out fróm són.
The cár will be out fróm in a mómént.

After describing the characteristic rhythmic pattern of English as one based on primary stresses, between which are squeezed varying numbers of lesser stresses, Smalley\(^{(142)}\) directed the student to notice how, on pre-recorded exercises, the syllables were rushed or spaced out to fill the space between beats and achieve even timing.

Part of a set follows.
The program was boring.
The program was not very boring.
The program we saw was very boring.
The program we saw was not very boring.

The authors of another text on English for non-native speakers (33) first advise the student that, though stress does not occur at absolutely regular intervals in everyday speech, it may be quite regular both in spontaneous utterance and in studied speech. A few pages farther on, nevertheless, the student is told that when "rhythm units are combined into sentences, primary stresses recur at practically equal intervals. As a result, several unstressed words plus a stressed one may take no longer to say than a stressed one." In the exercises following that observation, the student is asked to notice that, though the sentences in a set contain different numbers of syllables, they all contain an equal number of stressed syllables; therefore, they all require "practically" the same time to say. A few pairs of sentences of "practically" the same duration are

Learning takes work.
Getting A's requires concentration.

I want John.
They don't know how to cooperate.

John plays jazz.
His brother has played in big orchestras.

Take this along.
He's taken the rolls out of the oven.

Many foreign language textbooks for English speakers contrast the isochronic rhythm of English with a different type of rhythm in the target language. Two examples are Principes de phonétique française, in which Delattre (30) states that English rhythm is produced by the return of strong syllables at more or less equal intervals; and
Introducción a una comparación fonológica del español y del inglés, in which Cárdenas\(^{(19)}\) says that the duration of these sentences,

\begin{align*}
\text{The girl is pretty.} \\
\text{The girl isn't pretty.} \\
\text{The girl isn't very pretty.}
\end{align*}

would be the same, if the sentences were spoken, because the words between girl and pretty are reduced or lengthened to occupy the same time.

A few contemporary writers have questioned the theory postulated by Classe and by Pike. Interested primarily in the physiological bases of speech, Stetson\(^{(145)}\) observed that "the free rhythm of speech both in verse and prose is constituted not of like intervals in time, but of LIKE MOVEMENTS whose durations may vary by fifty per cent." De Groot\(^{(29)}\) called attention to the aesthetic significance of interrupted isochronism, in which the rhythmical effect becomes more distinct if not all the beats, but only two or three, are presented at equal intervals. Schramm\(^{(130)}\) found that experimentally he could not confirm any physical correlates of perceived rhythm.

After a survey of opinions on isochronism and a careful study of the instrumental measurements made by Classe, Baum\(^{(6)}\) was unconvinced about the isochronic basis of English prose rhythm and adopted a moderate position. He concluded that "strict isochronism is an ideal; by rigorous mechanical measurement it hardly exists except by accident or under rare circumstances." He added, however, that despite the existence of some arrhythmic prose, rhythmic prose does exist, but free from the demands of theoretical equality of measure. Since the rhythmic principle of equal or similar units can not be
denied, these units must be ascertained by other criteria which may be
either acoustic or expressional values, or certain rhetorical or syn-
tactic arrangements.

The second systematic attempt to investigate the timing of Eng-
lish prose rhythm by means of instrumental measurements concludes
with these words: "In brief, we did not find isochronism in our limited
data and therefore cannot say that there is isochronism in English
\(^{138}\)." Shen and Peterson, aided by Martin Joos in the analysis of the
data, spectrographed samples from prose readings by three linguists,
A. A. Hill, G. L. Trager, and W. F. Twaddell. The materials read
consisted of excerpts from works written by the linguists themselves,
as well as a set of sentences framed by K. L. Pike to illustrate iso-
chronism and recorded by him. The three linguists, recording at dif-
ferent times and places, did not know the purpose of the recordings.
Joos marked primary and secondary stresses as well as clause termi-
nals (according to the Trager and Smith system of phonemic analysis
of English), and pauses, on phonemic transcriptions of the recordings.
Part of the material had previously been transcribed into phonemics,
with primary stresses and terminal junctures already indicated by the
authors; in other parts terminals and pitches had been marked, "so
that primaries could be predicted." All stretches of speech with a
pause "more than just a terminal contour" between stresses were eli-
minated. The remaining stretches were measured from the end of a
consonant preceding one stressed vowel to the end of the consonant pre-
ceding the following stressed vowel.

Measurements, in seconds, of the time-lapse in 167 samples of
speech extending between two primary stresses and containing a clause terminal and open juncture, were compared. The frequency of occurrence of each time-lapse was also noted. Inspection of the tables showed a wide range of variation in the duration of the stretches for each subject: 0.41-1.82 sec., 0.38-2.5 sec., 0.55-3.61 sec. Finding that the evidence at this point was against isochronism, Shen and Peterson then measured and analyzed samples of intervals between primary and secondary /*/ stresses, and between secondary and secondary. As stipulated in the Trager and Smith system of phonemic analysis, none of these stretches might contain clause terminals. Joos condensed all the data (406 samples) in a table of distributions, on which each sample was represented by a dot. He disregarded three very short intervals (less than 19 cs.) on the basis of possible error, and 41 (10% of the total) extra-long intervals (above 82 cs.) on the basis that they were "longer than Pike's longest interval." Inspection of the intervals from 19 cs. to 81 cs. showed no clear picture; at most, there was some clustering at certain places but, preponderantly, too much scatter. From this evidence, Shen and Peterson concluded that "the postulated equality of various kinds of time intervals, subsumed under the term isochronism was not borne out."

The preceding summary includes all available studies touching on the question of isochronism in English prose. Apparently, no study previous to this one has been made of free, unrehearsed speech by homogeneous subjects whose recordings were analyzed without preconceived notions of where the stresses, pitch changes, transitions, or pauses might occur.
Restatement of the Problem

As applied by some linguists to English speech, the theory of isochronic rhythm specifies that the time-intervals between strongly stressed syllables are equal or approximately equal regardless of the number of unstressed syllables between major stresses (22, 62, 70, 90, 120, 168). The second part of this statement presupposes that there cannot be any direct relation between the duration of intervals and the number of syllables they contain, since the duration of the first interval in an isochronous rhythmic series must necessarily determine the duration of subsequent intervals.

The main purpose of this investigation is to test the second part of the theory: specifically, that in the free speech of a given subject, the duration of intervals between dominant stresses is independent of the number of syllables the intervals contain. If a direct relation between the duration of intervals and the number of syllables they contain should be found to exist, it would be necessary either to modify the theory or to discard it.

A practical application of the theory of isochronic prose rhythm frequently found in textbooks on the pronunciation of English is that the total duration of utterances containing the same number of primary stresses is equal, regardless of the number of syllables in the utterances and regardless of the segments preceding the first strong stress and following the last. In other words, it has been assumed that, for any given speaker, the number of primary stresses in an utterance is a coefficient of its duration (33, 119, 125, 142). To test the validity of this assertion is a secondary purpose of this investigation.
CHAPTER II. PROCEDURE

The experimental procedure adopted for this study can be summarized as follows:

1. A number of subjects representative of the North Midland dialect of American English recorded ten minutes of impromptu speech and two-minute readings of prepared material.

2. Orthographic transcriptions of the free speech were marked off into uninterrupted stretches, delimited by primary stresses.

3. Two hundred and two stretches containing different numbers of syllables were selected at random and spectrographed.

4. Strongly stressed syllables were identified on the spectrograms and the intervals between them measured in centiseconds.

5. Syllables contained between the onset of one strong stress and the onset of the next strong stress were identified and counted.

6. The duration of each sample and the number of syllables in it were plotted on individual frequency distributions and on a combined frequency distribution.

7. The data were analyzed for evidence of linear and
curvilinear relations.

8. Those framed utterances which had been read with the same number of strong stresses by several subjects were spectrographed, and their total duration was measured.

9. The observed duration of these utterances was compared with their predicted duration. Predicted durations were estimated on a basis of theoretical isochronic rhythm and on a basis of the average length of syllables in the shorter utterance.

**Validation of the Investigator's Judgment of Stress**

The initial task in this study required that primary stress points be identified in a corpus comprising over seventy-two minutes of recorded speech. Because this task would have been too time-consuming for a panel of expert observers, the attempt was made to ascertain whether the investigator could be validated as competent in the assignment of stress points. To determine this, the following procedure was adopted:

1. A validating panel consisting of five observers, three of them professors of speech and of phonetics, two of them graduate students with some experience in teaching phonetics, were asked to identify the dominant stresses in a five-minute recording of extemporaneous speech, which had been taped by a subject unknown to them.

2. The listeners were instructed to mark the "dominant
(prominent) stresses" they heard in parts of the recording that the investigator had previously selected and transcribed into conventional English orthography. These sections contained stretches of speech of various lengths, uninterruptible by noticeable pauses or hesitations.

3. Thirty-one samples of stretches of speech containing from one to ten syllables between the beginning of one stressed syllable and the beginning of the next stressed syllable were chosen from the marked passage at random.

4. On a basis of each listener's reaction to a given sample, a consensus, including the judgment of the investigator, was determined for each sample.

5. The percentage of agreement between each listener and the consensus was calculated on a basis of stress identifications in 27 samples, four samples having been eliminated because of evenly divided decisions. The results of the comparison are shown below.
### TABLE I

Percentage of Agreement between Listeners and Consensus

<table>
<thead>
<tr>
<th>Listener</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listener A</td>
<td>81%</td>
</tr>
<tr>
<td>Listener B</td>
<td>87%</td>
</tr>
<tr>
<td>Listener C</td>
<td>59%</td>
</tr>
<tr>
<td>Listener D</td>
<td>100%</td>
</tr>
<tr>
<td>Listener E</td>
<td>37%</td>
</tr>
<tr>
<td>Investigator</td>
<td>88%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>72%</strong></td>
</tr>
</tbody>
</table>

Inspection of the data presented above shows that the investigator achieved a higher level of agreement with the consensus than the average agreement, 72%, of the six observers. A comparison of the percentage of agreement between each listener and the consensus, in terms of rank order, shows that the investigator obtained the second highest percentage of agreement. On the basis of these results, the assumption was made that the investigator was able to make valid judgments concerning the location of stressed syllables in free speech.

Several factors may have contributed to the relatively low mean agreement among the observers:

1. the general problem of reaching agreement in a subjective analysis of speech;
2. the lack of an exact definition of the term "dominant (primary) stress";
3. the requirement that listeners agree on both stress
points delimiting a stretch of speech;
4. the fast and irregular tempo of impromptu speech, which makes the evaluation of degrees of stress more difficult;
5. the tendency of the listener to hear stress where he expects to hear it.

When all the possible sources of disagreement in making subjective judgments of stress were considered, the judgment shown by the investigator in this test was viewed as acceptable.

**Procedure Followed in Obtaining Corpus**

The procedure employed in obtaining the corpus used in the main part of this study is described in detail below.

Six informants (hereafter referred to as Subject 1, Subject 2, and so forth), all male college students in their early twenties, were chosen. Since all of them were born and raised in Akron, Ohio, and since all speak a similar form of the North Midland dialect of Standard American English, they may be considered linguistically unprejudiced. The informants, who were not instructed as to the nature of this study, talked spontaneously for ten minutes on any topic of their choice. (The topics included a summer in England, an engagement and forthcoming marriage, a radio workshop, golf, the ministry, and educational television.) The informants also read some sets of utterances, purported examples of isochronic rhythm extracted from various publications. Their speech was recorded on magnetic tape, at 7-1/2 inches per second, in a sound-treated recording room at the University of Akron.
An Ampex 350 tape-recorder and a high quality microphone were used.

**Procedure Followed in Obtaining Data on Impromptu Speech**

Each subject's recording of impromptu speech was transcribed into conventional orthography, and all stresses, major pitch changes, and pauses were marked. All stretches of continuous speech, enclosed by primary stresses and devoid of noticeable pauses, were grouped according to the number of syllables contained between the beginning of one primary stress and the beginning of the next primary stress. Each stretch contained any number of pitch changes and of close and open transitions. Only those open transitions which were manifested by acoustic characteristics of marginal sounds and not by pause were included\(^9\). The total inventory of free speech consisted of 1209 samples, each containing between one and eleven syllables.

For each speaker four samples were selected at random from all groups of stretches containing from one to six syllables; when fewer than five samples were available, however, all samples in the group were taken. As will be seen in the figures presented in Chapter III, Subjects 2 and 4 had only three samples of 7-syllable stretches; Subject 1 had only three samples, and Subject 3 no samples, of 8-syllable stretches; all subjects had fewer than four samples of 9- and 10-syllable stretches; and only Subjects 4 and 6 had an 11-syllable stretch. Absolute randomness was also impaired later by the fact that, upon attempting to measure the exact time-lapse of some stretches, an accurate identification of the point of demarcation could not be made. In such cases the next sample indicated on the table of random
sampling numbers was taken. A total of 202 samples of impromptu speech was then transferred to a Kay Electric Sonagraphe in the Speech Communication Laboratory at Louisiana State University, and a spectrogram (Sonagram) made of each.

**Spectrographic Analysis of the Data**

Briefly, spectrographic analysis of speech involves transforming the acoustic patterns of speech into visual patterns. The techniques of spectrographic analysis, developed during the last twenty years, are explained by Joos\(^{(83)}\), Barney and Dunn\(^{(5)}\), Hockett\(^{(71)}\), Potter, Kopp, and Green\(^{(124)}\), Halle\(^{(61)}\), and others\(^{(44, 61, 87)}\). Up to the present, four types of analysis have been found useful: broad-band analysis, narrow-band analysis, amplitude sections, and continuous amplitude displays. A general description of each follows.

1. Broad-band analysis. In this type of analysis an analyzing filter of 300 cps band-width scans the acoustic spectrum. The resulting spectrogram displays the formant structure of voiced sounds, the energy concentrations of voiceless sounds, and a time pattern of changes in frequency dimensions from 80 cps to 8000 cps. Broad-band spectrograms permit the study of the rate of change in the phonetic quality of speech sounds, against a time maximum of 2.2 or 2.4 seconds on each spectrogram. Broad-band analyses also display relative amplitudes grossly by means of lighter or darker markings. Figures 1 to 4 illustrate this type of analysis.
2. Narrow-band analysis. In this type of analysis a narrow (80 cps) analyzing filter is used to obtain a display of harmonic structure on the time-frequency plane of the stream of speech. Figure 5 illustrates narrow-band analysis.

3. Amplitude sections. These present the harmonic structure of a very brief segment, displaying the intensity of each harmonic and providing an opportunity to study closely the internal structure of the formants.

4. Continuous amplitude display. By means of a special circuit, it is possible to represent the overall intensity of a speech-wave as a continuous function of time. Thus, variations in amplitude from one cycle to another can be observed.

For the purpose of this study, broad-band spectrograms in most cases revealed enough significant clues for the identification of syllabic nuclei and of strong stress. In a few cases, however, it was necessary to use additional narrow-band (80 cps) analyses as an aid in the identification of the boundaries of stressed syllables. (See Figures 4 and 5.) No amplitude sections or continuous amplitude displays were used.

The sounds on each spectrogram were identified according to the established practice set forth by Potter, Kopp, and Green (124), Joos (83), Delattre, Liberman, and Cooper (31), Lehiste (93), and others (44, 61, 77, 87). In general terms, spectrographic patterns of speech sounds are differentiated as follows:
1. Duration is shown by the length of the pattern on the horizontal axis. The maximum duration on the spectrograms used was 2.2 seconds.

2. Frequency is represented on the vertical axis. On the spectrograms used in this study it ranged from 80 cps to 8000 cps.

3. Vowels, semi-vowels, glides, and nasals are portrayed by the number and position of resonance bars (dark horizontal bands), especially bars 1 and 2.

4. Consonants in general are distinguished by reduction of energy, and in particular by the different bending effects they have on the resonance bars of adjacent vowels.

5. Voiceless fricative continuants are portrayed by random vertical striations of different patterns, locations, and amplitudes.

6. Voiced fricative continuants are portrayed by vertical striations similar to those of their voiceless counterparts, but crossed by resonance bars and by a voice bar (a narrow resonance bar on the base line).

7. Voiceless stop-sounds are identified by a blank space (stop-gap) followed or preceded by a vertical line (glottal spike) and by random vertical striations.

8. Voiced stop-sounds are portrayed like their voiceless counterparts but with the addition of a voice bar on the base line and of resonance bars extending through the vertical striations.
9. Liquids are portrayed by a combination of the formant structure of vowels, additional formants in the lower frequencies, and, in most cases, evidence of the reduced energy characteristic of consonants.

10. Internal open transitions are identified by certain phenomena as the type of release of stop-sounds, glottal stops, increased duration, changes in intensity pattern, unvoicing of normally voiced sounds or vice-versa, and laryngealization (slow, irregular, vocal-fold activity with reduced energy)\(^{(93)}\).

In order to measure the duration of stretches of speech from stressed syllable to stressed syllable, it was necessary to select a precise point of demarcation. A point dividing the onset of a stressed vowel from the consonant sound preceding it or from silence was chosen as the demarcation point for the following reasons:

1. In rhythmic grouping the accent tends to be placed on the first or last element of a group\(^{(6, 47, 66, 116, 133, 144, 153, 166)}\).

2. Most authorities agree that in English the indispensable part of a stressed syllable is a vowel. (See Chapter I.)

3. The onset of stress is most often perceived as concurrent with the onset of the vowel\(^{(27, 28, 49, 53, 67, 77, 99, 142, 150, 153, 159)}\).

4. A preceding consonant seems to have less effect on a vowel than a following consonant\(^{(118)}\).

5. This is the same demarcation point used by Classe\(^{(22)}\).
Fig. 1. Segmentation of adjacent two-syllable (30 cs) and six-syllable (87 cs) stretches spoken by Subject 4.
Fig. 2. Segmentation of three-syllable stretch measuring 51cs
Fig. 3. A seven-syllable stretch (120 cs) adjacent to a three-syllable stretch spoken by Subject 3.
Fig. 5. Narrow band spectrogram
One-syllable stretch shown in Fig. 4
Duration - 22 cs
and by Shen and Peterson\textsuperscript{(138)} in their studies of isochronism.

Stressed vowels, previously identified subjectively, were then identified acoustically on the spectrograms by relative darkness and relative duration of the spectrum pattern, and, in some cases, by indications of major pitch change in the resonance bars. At times, when the identification of strong stress was debatable, the peaks of output of energy indicated on the VU-Meter attached to the spectrograph were used as corroboratory evidence. Demarcation points at the onset of each strongly stressed vowel were then determined on all spectrograms of samples of free speech. (See Figs. 1 to 5.) The stretches between these points were then measured in centiseconds by means of a plastic overlay. The accuracy of the calibrator was checked against a spectrogram of a 20 cps pure tone, on which the distance between spikes represented 5-centisecond intervals.

The number of syllables in each sample was determined on the basis of the number of syllabic centers (vowels, diphthongs, and the so-called syllabic consonants $\left[\text{m} \right], \left[\text{n} \right], \left[\text{x} \right], \left[\text{l} \right]$) observed. The latter were considered as syllabic centers and counted as syllables only when their spectra warranted it. The data for each subject were then plotted on Figures 6 to 11, and the combined data on Figure 12.

Inspection of the graphs suggesting the existence of a relationship between the number of syllables and the duration of each stretch, the data were then processed in the Computer Research Center at Louisiana State University in order to investigate the nature of the relationship. The findings are presented in Chapter III.
Procedure Followed in Obtaining Data on Structured Speech

The recordings of sets of framed utterances given by various authors (33, 70, 119, 125, 142) as illustrations of isochronic rhythm were handled in a different fashion. Seven sets of utterances, containing a total of 28 utterances, were read by the informants and recorded on tape. Since the informants were instructed merely to "read naturally," there was wide variation in the readings. Those sets read with disparate numbers of stresses in the different utterances composing the set were, of course, discarded. The remaining sets, read with the same number of strong stresses by several subjects, were selected for comparison.

The 71 samples to be compared were submitted to spectrographic analysis, and their duration was measured from the beginning of speech to its cessation, since the sentences were presumed by their authors to be equal or approximately equal in duration because of the effect of isochronism. The observed measurements of total duration were tabled, and two types of predictions of their duration were made and compared.

For the four sets of three similar sentences read with the same number of strong stresses, such as

The program wasn't boring.
The program was very boring.
The program wasn't very boring (142).

predictions of the probable duration of the second and third sentences in each set were made on the basis of theoretical isochronic rhythm. Thus, it was assumed that, because of the effect of isochronism, the total duration of successive sentences in the same set would be equal
to the duration of the first sentence. The accuracy of these predictions was compared to that of predictions made on a basis of the average duration of syllables in the first sentence of each set, on the assumption that there might be a linear relation between the number of syllables in an utterance and the duration of the utterance.

For the three pairs of dissimilar utterances, such as

Look! Look! Look!
That's all I ask, that you look\(^{33}\).

presumed by their authors to be of equal duration because of the tendency toward isochronic rhythm, predictions of the duration of the second utterance were made in the same fashion as those explained above.

The utterances, their observed duration, the number of syllables and the location of strong stresses in each are shown on Table V. A comparison of the accuracy of prediction on a basis of the theory of isochronic rhythm and on a basis of a linear relation are discussed in the following chapter.
Fig. 6 Duration of 33 Samples of Stress Groups in Impromptu Speech. Subject 1.

Number of syllables
- Sample
- Mean average
Fig. 7 Duration of 31 Samples of Stress Groups in Impromptu Speech. Subject 2.

Number of syllables
- Sample
- Mean average
Fig. 8 Duration of 30 Samples of Stress Groups in Impromptu Speech. Subject 3.
Fig. 9 Duration of 32 Samples of Stress Groups in Impromptu Speech. Subject 4.

Number of syllables
- Sample
- Mean average
Fig. 10 Duration of 33 Samples of Stress Groups in Impromptu Speech. Subject 5.
Fig. 11 Duration of 35 Samples of Stress Groups in Impromptu Speech. Subject 6.

---

Sample

Mean average
Fig. 12 Combined Results. 180 Samples.

Number of syllables

Duration in centiseconds

- Sample
- Mean average
CHAPTER III. RESULTS

Interpretation of Data on Impromptu Speech

The random samples of stretches of uninterrupted speech between primary stresses uttered spontaneously by each subject are plotted on Figures 6 to 11, the combined data on Figure 12. On each graph the independent variable, i.e. the number of syllables in each stretch of speech, is shown on the horizontal axis. The dependent variable, duration, measured in centiseconds, is shown on the vertical axis. Only stretches containing from one to eight syllables are shown on Figure 12, the number of stretches longer than these not being sufficient for a predictive study.

Inspection of the graphs shows two possibilities of relation between the two variables:

1. a probable linear regression for all relationships tested;
2. a possible curvilinear relation in the extremes of the function for Subjects 1, 2, and 4.

The possibility of a curvilinear effect in any subject or in the combined subjects can be tested by applying the quadratic equation

\[ \hat{y} = b_0 + b_1 x_1 + b_2 x^2 \]

in which

\[ \hat{y} \] represents the predicted time interval for a given number of syllables,

\[ b_0 \] represents the \( y \)-intercept,
\[ b_1 \] represents the linear effect, and
\[ b_2 \] represents the curvilinear effect.

The values for each factor in the equation, obtained from the initial processing of the data, are shown in the table below.

**TABLE II**

Preliminary Values of Linear and Curvilinear Effects in the Relation between Number of Syllables and Duration

<table>
<thead>
<tr>
<th>Subject</th>
<th>( b_0 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>12.00</td>
<td>13.99</td>
<td>0.16</td>
</tr>
<tr>
<td>Subject 2</td>
<td>21.62</td>
<td>8.55</td>
<td>0.65</td>
</tr>
<tr>
<td>Subject 3</td>
<td>8.37</td>
<td>14.44</td>
<td>0.09</td>
</tr>
<tr>
<td>Subject 4</td>
<td>27.14</td>
<td>6.18</td>
<td>0.57</td>
</tr>
<tr>
<td>Subject 5</td>
<td>20.01</td>
<td>8.74</td>
<td>0.12</td>
</tr>
<tr>
<td>Subject 6</td>
<td>8.46</td>
<td>14.21</td>
<td>0.31</td>
</tr>
<tr>
<td>Combined subjects</td>
<td>14.19</td>
<td>12.46</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Since a comparison of these values strongly suggests the possibility that \( b_1 \) may be significant while \( b_2 \) may not be significant, a t-test is now indicated as a definitive test of the significance of \( b_1 \) and \( b_2 \), that is, as a determination of whether \( b_1 \) and \( b_2 \) are different from 0.

The values of \( b_1 \) and \( b_2 \) obtained by running a t-test are shown in the table that follows.
TABLE III
Significance of Linear and Curvilinear Effects in the Relation between Number of Syllables and Duration

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>4.41</td>
<td>.50</td>
</tr>
<tr>
<td>Subject 2</td>
<td>2.00</td>
<td>.14</td>
</tr>
<tr>
<td>Subject 3</td>
<td>5.49</td>
<td>.34</td>
</tr>
<tr>
<td>Subject 4</td>
<td>1.58</td>
<td>.14</td>
</tr>
<tr>
<td>Subject 5</td>
<td>2.34</td>
<td>.33</td>
</tr>
<tr>
<td>Subject 6</td>
<td>4.21</td>
<td>.95</td>
</tr>
<tr>
<td>Combined subjects</td>
<td>7.49</td>
<td>.17</td>
</tr>
</tbody>
</table>

From a comparison of these values the conclusion can be drawn that the value of $b_1$ (the linear effect) is significant, and that the value of $b_2$ is not significant for any subject or for the combined subjects. Therefore, a curvilinear effect can be considered negligible.

To test the first possibility, a linear regression, a first degree equation in $x$ may be used. This equation will yield a coefficient of correlation; that is, a single number that shows to what extent two measurables are related. In this particular case, a coefficient of correlation will show to what extent the variations in time are associated with the variations in number of syllables. In theory, a straight line relation should have a correlation coefficient of 1.; however, a correlation coefficient of approximately 1., such as .80, strongly suggests a straight line relation between two variables.

Application of the general equation of a straight line,

$$\hat{y} = b_0 + b_1 x$$

to the data in this investigation (202 paired observations, 199 degrees of freedom) yields the values presented in Table IV.
TABLE IV

Coefficients of Correlation between Number of Syllables and Duration

<table>
<thead>
<tr>
<th>Subject</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>.948</td>
</tr>
<tr>
<td>Subject 2</td>
<td>.940</td>
</tr>
<tr>
<td>Subject 3</td>
<td>.975</td>
</tr>
<tr>
<td>Subject 4</td>
<td>.918</td>
</tr>
<tr>
<td>Subject 5</td>
<td>.906</td>
</tr>
<tr>
<td>Subject 6</td>
<td>.930</td>
</tr>
<tr>
<td>Combined subjects</td>
<td>.916</td>
</tr>
</tbody>
</table>

The coefficients above, in every case greater than .90, show a high positive correlation between the number of syllables in a stretch and the duration of the stretch. They establish with little doubt that, for any subject similar to those investigated, the duration of an uninterrupted stretch of impromptu speech contained between primary stresses can be predicted on the basis of the number of syllable in the stretch. This conclusion does not support, but rather negates, the theory of isochronic rhythm, according to which there can be no close relation between the number of syllables in a stretch and its duration.

Interpretation of Data on Structured Speech

As stated in Chapter I, a secondary purpose of this study is to test the validity of an assumption made by the authors of several textbooks, in effect, that even though the utterances in certain sets of examples contain a different number of syllables, they are equal or approximately equal in total duration because they have the same number of primary stresses. As may be seen in Table V below, Sets I to IV consist of three phonetically and grammatically similar sentences;
TABLE V

Observed Duration of Framed Utterances

<table>
<thead>
<tr>
<th>Sets of Utterances</th>
<th>Duration, in cs., for each Subject</th>
<th>Mean No. of Aver. Sylls.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S. 1</td>
<td>S. 2</td>
</tr>
<tr>
<td>I. Boys need money.</td>
<td>110</td>
<td>115</td>
</tr>
<tr>
<td>The boys will need some money.</td>
<td>112</td>
<td>113</td>
</tr>
<tr>
<td>The boys will be needing some money.(125).</td>
<td>112</td>
<td>113</td>
</tr>
<tr>
<td>II. English is easy.</td>
<td>112</td>
<td>152</td>
</tr>
<tr>
<td>English is very easy.</td>
<td>112</td>
<td>152</td>
</tr>
<tr>
<td>The English lesson is very easy.(119).</td>
<td>112</td>
<td>152</td>
</tr>
<tr>
<td>III. The speech ended.*</td>
<td>112</td>
<td>132</td>
</tr>
<tr>
<td>The speech was boring.*</td>
<td>112</td>
<td>132</td>
</tr>
<tr>
<td>The program was boring.(142).</td>
<td>112</td>
<td>132</td>
</tr>
<tr>
<td>IV. The program wasn't boring.</td>
<td>170</td>
<td>131</td>
</tr>
<tr>
<td>The program was very boring.</td>
<td>162</td>
<td>156</td>
</tr>
<tr>
<td>The program wasn't very boring.(142).</td>
<td>176</td>
<td>164</td>
</tr>
<tr>
<td>V. Find a board for me.</td>
<td>170</td>
<td>131</td>
</tr>
<tr>
<td>Interpret this poetry for me.(70).</td>
<td>162</td>
<td>156</td>
</tr>
<tr>
<td>VI. Great progress is made daily.</td>
<td>176</td>
<td>164</td>
</tr>
<tr>
<td>The boy is interested in enlarging his vocabulary.(125).</td>
<td>174</td>
<td>136</td>
</tr>
<tr>
<td>VII. Look! Look! Look!</td>
<td>177</td>
<td>146</td>
</tr>
<tr>
<td>That's all I ask, that you look.(33).</td>
<td>133</td>
<td>161</td>
</tr>
</tbody>
</table>

*Framed by investigator to extend a group by Smalley.(142).
Sets V, VI, and VII are pairs of utterances that differ phonetically, grammatically, or both. The total duration of those utterances in each set which were read with the same number of strong stresses and with no pauses (except the utterances in Set VII) was measured from the first display of speech until the last. These measurements, together with the number of syllables and the observed duration of each utterance, as read by different subjects, are shown in Table V.

Inspection of Table V shows that in every case except two (the reading of Set VII by Subject 1 and of Set IV by Subject 2) the second utterance had a longer overall duration than the first, and the third a longer duration than the second. Moreover, the mean average for every set shows the same result, of course. Though the data obtained from these readings are contaminated by the inclusion of syllables preceding the first and following the last strong stress, the conclusion must be reached that in framed sets such as these, isochronic rhythm is not evident.

A practical test of the empirical statement found in many textbooks for teaching English as a foreign language, that the duration of utterances in American English can be considered in terms of an isochronic theory, may be made by comparing the accuracy of predicting the duration of paired utterances in terms of the isochronic theory with the accuracy of predicting duration in terms of a hypothetical linear relation between the duration of an utterance and the number of syllables it contains.

Since, according to the theory of isochronism, each of the paired utterances in Table V should occupy the same time, the accuracy of this theory in terms of its predictive value can be calculated by
dividing the difference between observed time \( t_o \) and expected time \( t_e \) by the expected time. The resulting figure would represent the percentage of error in the calculation. By subtracting it from 100, the percentage of accuracy can be derived. The basic formula to be applied is

\[
\text{Percentage of prediction accuracy} = 100 - \frac{t_o - t_e}{t_e}
\]

The percentage of prediction accuracy in terms of a hypothesis of linearity can be obtained by using the same formula as above. However, the expected duration \( t_e \) of the second of two paired utterances is obtained in this case by multiplying the average duration of syllables in the observed utterance by the number of syllables in the utterance the duration of which is to be predicted.

Comparisons of the percentages of prediction accuracy found in the utterances shown in Table V are given in the tables that follow.

**TABLE VI**

Comparison of the Number of Accurate Predictions of Duration in Terms of the Theory of Isochronism and in Terms of a Hypothesis of Linearity

<table>
<thead>
<tr>
<th>Number of Predictions</th>
<th>Isochronic More Accurate</th>
<th>Linear More Accurate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similar utterances</td>
<td>34</td>
<td>7 (20%)</td>
</tr>
<tr>
<td>Dissimilar utterances</td>
<td>10</td>
<td>3 (30%)</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>10 (22%)</td>
</tr>
</tbody>
</table>
**TABLE VII**

Comparison of the Mean Average Percentage of Prediction Accuracy in Terms of the Theory of Isochronism and in Terms of a Hypothesis of Linearity

<table>
<thead>
<tr>
<th>Sets of Similar Utterances</th>
<th>Isochronic Prediction Accuracy</th>
<th>Linear Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set I</td>
<td>49%</td>
<td>55%</td>
</tr>
<tr>
<td>Set II</td>
<td>38%</td>
<td>97%</td>
</tr>
<tr>
<td>Set III</td>
<td>70%</td>
<td>93%</td>
</tr>
<tr>
<td>Set IV</td>
<td>87%</td>
<td>93%</td>
</tr>
<tr>
<td>Average</td>
<td>61%</td>
<td>85%</td>
</tr>
</tbody>
</table>

Sets of Dissimilar Utterances

<table>
<thead>
<tr>
<th>Sets</th>
<th>Isochronic Prediction Accuracy</th>
<th>Linear Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set V</td>
<td>62%</td>
<td>78%</td>
</tr>
<tr>
<td>Set VI</td>
<td>68%</td>
<td>67%</td>
</tr>
<tr>
<td>Set VII</td>
<td>64%</td>
<td>54%</td>
</tr>
<tr>
<td>Average</td>
<td>64%</td>
<td>65%</td>
</tr>
</tbody>
</table>

Overall Average 62% 77%

Analysis of the accuracy of prediction in terms of subjects shows that prediction on a basis of a linear relation is consistently higher for all subjects except Subject 2, in whose case it is lower in two out of four predictions made for that subject. In Subjects 1, 4, 5, and 6 a linear prediction is higher in all comparisons except one for each subject.

The preceding comparisons are, of course, crude tests; but they indicate that the assumption whereby the strong stresses in an utterance may be taken as coefficients of its duration does not hold up under preliminary analysis. In fact, even in grammatically and phonetically similar utterances, in which, according to Classe\(^{(22)}\), the circumstances are most favorable for isochronism, predictions based on this
theory are less accurate than those based on a hypothesis of linearity.

The fact that predictive accuracy based on a theory of linearity is not more than 77% accurate may be due to unanalyzed factors, such as differences in tempo. More probably it is due to one or both of the following factors:

1. the inclusion, in some cases, of the duration of one or two unstressed syllables occurring before the onset of strong stress in one of two paired utterances;
2. the observed decrease in mean average duration of syllables in multisyllabic stress groups.

The decrease in mean average duration of syllables in multisyllabic stress groups is a commonly accepted observation often cited as an essential element in achieving isochronic rhythm. The data provided by the samples of free speech in this study confirm this observation. Inspection of Table VIII below shows that the average duration of syllables in stress groups decreases as the number of syllables in the group increases. The decrease, however, is neither directly proportional to the increase in number of syllables nor large enough to equalize the duration of groups containing an unequal number of syllables.

Moreover, though inspection of Table VIII confirms the well-known fact that stressed monosyllables are proportionally longer than syllables in polysyllabics, it makes apparent, nevertheless, that the mean average duration of 2- and 3-syllable stress groups is almost the same, differing by only 0.6 cs. (The equal duration of syllables in the more common 2-syllable stress groups has long been observed by Daniel Jones(79).) Inspection of the table also shows that the average duration of
syllables in longer stress groups, containing from 4 to 8 syllables, varies within a maximum of only 1.5 cs. It may be reasonably assumed, then, that the shortened duration of syllables in multisyllabic stress groups is a negative factor affecting a possible 1:1 ratio between an increase in number of syllables and an increase in duration.

### TABLE VIII

Average Duration of Syllables in 180 Samples of Stress Groups from the Free Speech of All Subjects

<table>
<thead>
<tr>
<th>Number of Syllables in Group</th>
<th>Mean Average Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.6 cs.</td>
</tr>
<tr>
<td>2</td>
<td>19.2 cs.</td>
</tr>
<tr>
<td>3</td>
<td>19.8 cs.</td>
</tr>
<tr>
<td>4</td>
<td>15.6 cs.</td>
</tr>
<tr>
<td>5</td>
<td>14.9 cs.</td>
</tr>
<tr>
<td>6</td>
<td>15.0 cs.</td>
</tr>
<tr>
<td>7</td>
<td>14.2 cs.</td>
</tr>
<tr>
<td>8</td>
<td>14.1 cs.</td>
</tr>
</tbody>
</table>

A cursory look at the duration of abutting stress groups found incidentally in the samples of free speech randomly selected for this study yields what may be further evidence against the isochronic theory. The disparity in the duration of adjacent stress groups presented graphically in Figures 13 and 14 indicates that, even in instances of exact conjunction of stress groups, there is not equal duration. Of the 16 pairs of abutting stress groups in Figure 13, none are equal. Two pairs, differing by 6 cs. or less, may be said to be approximately
equal. In both pairs, however, the two abutting stretches contain the same number of syllables. Figure 14 presents graphically the duration of 12 pairs of adjacent stress groups in the free speech of Subject 4, who showed the strongest tendency toward isochronic rhythm. Nevertheless, only 5 of the pairs are equal or approximately equal, and one of these contains the same number of syllables. The proportion of equal or approximately equal abutting stress groups for all subjects is presented in Table IX.

**TABLE IX**

<table>
<thead>
<tr>
<th>Subject</th>
<th>No. of Adjacent Pairs</th>
<th>No. of Pairs of Equal Duration</th>
<th>No. of Pairs of Approximately Equal Duration</th>
<th>No. of Equal or Approximately Equal Pairs Containing Same Number of Syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>2</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Percentage</td>
<td>3.5%</td>
<td>17.8%</td>
<td>41.6%</td>
<td></td>
</tr>
</tbody>
</table>

The figures presented in Table IX are, of course, raw data based on a very limited number of samples. They do suggest, however, the need for further research and statistical analysis. Needless to say, they have not been taken into consideration in the conclusions of this study, which are presented in Chapter IV.
Fig. 13 Duration of Adjacent Stress Groups in Free Speech. Subject 1.
Fig. 14 Duration of Adjacent Stress Groups in Free Speech. Subject 4
CHAPTER IV. CONCLUSIONS

Summary

The purpose of this study is to test the validity of the general belief that the rhythm of American English speech is isochronic, that is, a rhythm characterized by the recurrence of equal or approximately equal sequences of unstressed or weakly stressed syllables alternating with strongly stressed syllables which mark the rhythmic beat.

Two previous studies have attempted to test experimentally the theory of isochronic rhythm in English. In the first, published in 1939, André Classe(22) concluded that, under favorable circumstances, the rhythm of English speech is in effect isochronic; and that, even under unfavorable circumstances, a tendency toward isochronism is evident. In the second study, published in 1962, Yao Shen and Giles Peterson(138) found no evidence of isochronism in English rhythm. Both of these investigations were based on measurements of prose readings.

A secondary purpose of this study is to test the assertion made in some textbooks for the teaching of English as a foreign language that, because of the isochronous character of English rhythm, the number of primary stresses in an utterance determines its total duration, regardless of the number of syllables the utterance may contain. In effect, the assumption is that, for any given subject, the primary stresses in an utterance may be considered as coefficients of its
duration, so that two utterances containing the same number of strong stresses are presumed to be of equal duration regardless of the number of syllables in either utterance.

The corpus for the main part of this study consists of 1200 samples of stretches of speech obtained from ten-minute recordings of impromptu speech by six homogeneous subjects, all speakers of standard North Midland dialect of American English. The corpus for the secondary part of this study consists of recordings, by the same subjects, of 128 examples of isochronic rhythm taken from textbooks. All recordings were made with professional equipment.

The 200 samples of uninterrupted free speech by each subject were grouped according to the number of syllables they contained, ranging from one to eleven, and randomized. Four samples from each group, and all samples from groups containing fewer than five, were spectrographed. Syllabic centers and strong stresses were identified on each spectrogram, and the duration of each stretch was measured from the onset of one stressed vowel to the onset of the next stressed vowel. The duration of each stretch as well as the number of syllables in each stretch was plotted on a frequency distribution for each subject, each distribution containing from 30 to 35 paired measurements. After discarding stretches containing more than eight syllables because of their small number and because of possible error, the data for the 180 remaining stretches containing from 1 to 8 syllables were plotted on a combined frequency distribution.

Since inspection of all seven graphs showed a probable linear relation and a possible curvilinear relation between duration and number of syllables, the data were submitted to mathematical analysis for
these two effects. The analysis for each subject and for the combined subjects shows that, while the curvilinear effect is non-significant, the linear effect is significant in all cases. A high degree of correlation, in every case above .90, exists between the duration of and the number of syllables in a stretch of speech bounded at each end by the onset of strong stress.

The 128 recordings of framed utterances presented in various textbooks as examples of the effect of isochronic rhythm on duration were analyzed for similarity of rendition. Seventeen groups of three similar utterances each and 10 pairs of dissimilar utterances, totaling 61 samples, were found to have been spoken with the same number of strong stresses by three or more informants. These samples were spectrographed, their duration was measured from absolute beginning to absolute end, and the measurements were compared. Inspection of Table V shows that in 42 of the 44 comparisons the utterance containing more syllables has a greater observed duration than the utterance containing fewer syllables.

To test further the validity of the assumption that, because of the effect of isochronism, the duration of an utterance can be predicted on a basis of the number of strong syllables with which it is spoken rather than on a basis of the number of syllables it contains, predictions of the duration of these framed utterances were made in terms of the two theories. On a basis of the theory of isochronism, the predicted duration of the second of a pair of dissimilar utterances and of the second and third of a set of similar utterances, spoken with the same number of strong stresses, was assumed to be the same as that of the first utterance. On a basis of a theory of linear relation, the duration of a
second and of a third utterance was estimated to be the product of the number of syllables in it times the average duration of syllables in the first utterance.

Comparison of the accuracy of each type of prediction shows that

1. in terms of the number of predictions, those predictions based on a theory of linearity are more accurate in 77% of the 44 comparisons made;

2. in terms of percentage of predictive accuracy, those predictions based on a theory of linearity have an average accuracy of 75% as against 62% for those based on a theory of isochronic rhythm.

General Conclusions

Interpretation of the limited data in this study leads to the conclusion that isochronic rhythm, as defined in Chapter I, does not exist in the free conversation of speakers of North Midland dialect of American English. As postulated, the theory of accentual periodicity rests on the assumption that strong stresses recur regularly in the stream of speech regardless of the number of syllables encompassed by the stresses. Nevertheless, the physical measurements of time intervals between stresses presented in this study do not support such a theory. On the contrary, the evidence here presented contradicts the assumption by revealing the existence of a direct relationship between the number of syllables in stretches of speech delimited by strong stresses and the duration of those stretches. In fact, in all subjects studied the high degree of correlation between these two variables makes it
possible to set up a tentative hypothesis of linear relation between duration and number of syllables.

The hypothesis here presented is that the duration of stretches of speech delimited by the onset of strong stress increases as the number of syllables in the stretch increases. Though the increase is not directly proportional, since the mean average duration of syllables in multisyllabic stress groups, especially in those containing four or more syllables, tends to be smaller than that of syllables in shorter groups, it is clearly evident in all subjects. Furthermore, in none of the subjects does the average duration of individual syllables in multisyllabic stress groups containing a disparate number of syllables decrease enough to equalize the total duration of the groups.

Evidence provided by the comparison of the duration of framed utterances answers the secondary purpose of this investigation: to test the validity of the assertion that utterances containing an equal number of primary stresses but a disparate number of syllables are equal in total duration. The conclusion to be drawn from the measurements made in this study is that this assertion has no basis in fact, since in 95% of the comparisons the utterance containing the greater number of syllables also has a longer duration. The assumption, then, that because of the effect of isochronism the number of strong stresses in an utterance may serve as a coefficient of its duration, must be regarded as false. Moreover, the fact that in the majority of the comparisons the duration of the longer utterance could be predicted more accurately on a basis of average syllable length than on a basis of equal length of stress groups provides still more evidence against the theory of isochronic distribution of stresses and in favor of a
hypothesis of linearity.

**Suggestions for Further Studies**

Though it has been generally accepted that rhythm is an essential characteristic of speech, is it warranted to assume that free, unrehearsed, conversational speech is also rhythmical? Since all previous studies of the rhythm of speech have been based on oratory or on readings, in most cases readings of literary prose, it is not yet possible to assert that ordinary, informal speech is rhythmical. It seems that the most fundamental and hence most necessary investigation to be made on this subject would be one to determine whether or not listeners perceive a rhythmic impression from this type of speech. If impromptu speech should be found to be rhythmical, the next step would be to determine what factors constitute its rhythm.

Rhythm is, by general consent, a repetition or recurrence of some sort, involving some kind of time relation; if applied to speech, it must involve a recurrence of some sort among sounds. This recurrence, however, may involve not only units of duration or intensity, but also other acoustic factors such as changes in pitch, tempo, or sound quality, or even periods of silence. In addition, it is possible that a rhythmic impression may be conveyed by other types of expression units such as rhetorical or syntactic arrangements. It is also possible that the rhythmic quality of speech may be a complex phenomenon combining many factors. It would certainly be advisable to investigate the following possibilities:

1. that the beats in a rhythmic sequence may be supplied
not only by stresses but also by pitch changes or by pauses;

2. that the rhythmic impression may be conveyed merely by the alternation of stress and unstress without the constraint of periodicity;

3. that the rhythmic impression conveyed by speech may be solely or primarily a subjective phenomenon in which successive units are perceived in proportioned order without basis in objective measurements.

It would be advisable, of course, to conduct further investigations in the alleged isochronic nature of speech rhythm. A comparison of the duration of adjacent stretches in a rhythmic series, viewed cursorily in this study, merits statistical analysis. An investigation of the effect of hesitations, involuntary repetitions, curtailed rhythm units(119), non-terminal pauses, and juncture drawl (slow-down) on the alleged regularity of stress distribution may clarify or, on the other hand, invalidate the findings of this study. So may the investigation of phenomena such as syllables with double nuclei(122) and complex rhythm units containing two or more strong stresses(119), which perhaps act as disrupting factors in the postulated equality of stress groups.

The fact remains, nevertheless, that the evidence resulting from this study does not confirm the existence of isochronic rhythm in American speech.
BIBLIOGRAPHY


122. Pike, Kenneth L. "Practical Phonetics of Rhythm Waves," 

    of Vowels and Their Movements," *The Journal of the 
    Acoustical Society of America*, 20 (1948), 528-535.

124. __________, George A. Kopp, and Harriet C. Green, *Visible 

    151 pp.


127. Saussure, Ferdinand de. *Cours de linguistique générale*, 3d 

128. Schipper, J. *Grundriss der englischen Metrik*, Leipzig: 
    Braumüller, 1895. 404 pp.

129. Schramm, W. L. "The Acoustical Nature of Accent in American 

130. __________. "Approaches to the Study of English Verse," *Uni- 
    versity of Iowa Studies*, Series on Aims and Progress of 
    Research, No. 46, 1935. 82 pp.

131. Schubiger, Maria. *The Role of Intonation in Spoken English*, 

132. Scott, F. N. "The Scansion of Prose Rhythm," *Publications of 
    the Modern Language Association*, 20 (1905), 707-729.

133. Scott, John H. *Rhythmic Prose*, *University of Iowa Humanistic 

    (1940), 55-57.

135. __________. "An Experiment on Stress Perception," *Le Maître 
    Phonétique*, 3 (1939), 44-45.

136. Scripture, Edward W. *The Elements of Experimental Phonetics*, 

137. __________. *Researches in Experimental Phonetics*, Washington, 

138. Shen, Yao, and Giles G. Peterson. *Isochronism in English*, Stud- 
    ies in Linguistics, Occasional Papers No. 9, University of


## APPENDIX

### STRESS GROUPS REPRESENTED IN FIGURES 6-11

<table>
<thead>
<tr>
<th>Type of Group</th>
<th>DURATION IN CS.</th>
<th>Type of Group</th>
<th>DURATION IN CS.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-syllable</strong></td>
<td></td>
<td><strong>6-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>néxt fáll</td>
<td>26</td>
<td>pós(i)bly go on for a doctoralate</td>
<td>79</td>
</tr>
<tr>
<td>stáge túrnéd</td>
<td>29</td>
<td>appeared that I only had half cast party as I recall</td>
<td>83</td>
</tr>
<tr>
<td>léaped báck</td>
<td>34</td>
<td>almost knocked me off the pier</td>
<td>90</td>
</tr>
<tr>
<td>ódd jób스</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2-syllable</strong></td>
<td></td>
<td><strong>7-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>tháit or nót</td>
<td>27</td>
<td>probably ended up in a college</td>
<td>94</td>
</tr>
<tr>
<td>thís and dówn</td>
<td>29</td>
<td>nót for any association</td>
<td>94</td>
</tr>
<tr>
<td>póked with hóles</td>
<td>36</td>
<td>cábin on the edge of the stage</td>
<td>108</td>
</tr>
<tr>
<td>things that happened</td>
<td>38</td>
<td>smashed the door away with the axe</td>
<td>112</td>
</tr>
<tr>
<td><strong>3-syllable</strong></td>
<td></td>
<td><strong>8-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>fáirly goot cást</td>
<td>42</td>
<td>I don't know it must have been at least</td>
<td>92</td>
</tr>
<tr>
<td>cabóose and a cóal</td>
<td>44</td>
<td>on stage we were having a good time</td>
<td>122</td>
</tr>
<tr>
<td>hálf of an árm</td>
<td>48</td>
<td>blinded me I was looking right át</td>
<td>123</td>
</tr>
<tr>
<td>cóal cár and éverything</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4-syllable</strong></td>
<td></td>
<td><strong>9-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>dówn into a cóat</td>
<td>59</td>
<td>idea as to what I wanted to do</td>
<td>123</td>
</tr>
<tr>
<td>tíghtened up líke thís</td>
<td>62</td>
<td>área that I felt a littlé more cómfortable</td>
<td>124</td>
</tr>
<tr>
<td>take all sortes of ódd</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>túrnéd on thís húge spót</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5-syllable</strong></td>
<td></td>
<td><strong>10-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>í think it's prettí gúd</td>
<td>65</td>
<td>everybody in the audience was lágüing</td>
<td>98</td>
</tr>
<tr>
<td>Dóoditch was play(i)ng the víllain</td>
<td>80</td>
<td>I don't remember whether it was myself</td>
<td>110</td>
</tr>
<tr>
<td>áxe there to knock the dór</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>báck up on to the píer</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Group</td>
<td>Duration in cs.</td>
<td>Type of Group</td>
<td>Duration in cs.</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>1-syllable</strong></td>
<td></td>
<td><strong>7-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>bad after</td>
<td>20</td>
<td>práctica(lly) everything I do</td>
<td>103</td>
</tr>
<tr>
<td>back out</td>
<td>25</td>
<td>answering my problem at all</td>
<td>115</td>
</tr>
<tr>
<td>first one</td>
<td>36</td>
<td>competition against yourself</td>
<td>122</td>
</tr>
<tr>
<td>each shot</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2-syllable</strong></td>
<td></td>
<td><strong>8-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>mistake of my</td>
<td>34</td>
<td>Mógador is not a very difficult</td>
<td>92</td>
</tr>
<tr>
<td>excél at all</td>
<td>34</td>
<td>really going to master that game</td>
<td>140</td>
</tr>
<tr>
<td>understand the feeling</td>
<td>44</td>
<td>trying to better his previous score</td>
<td>144</td>
</tr>
<tr>
<td>succeed this year</td>
<td>51</td>
<td>game begins to get the best of you</td>
<td>146</td>
</tr>
<tr>
<td><strong>3-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gotten in shape</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>five hundred yards</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>didn't excél</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>being left-handed</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>goes into the golf</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>did have a good time</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>this was the mistake</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>left-handed in practically</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hole was about a hundred</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>position to play the ball</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>after about the third</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>this portion of the game</td>
<td>86</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>6-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>have to discipline yourself</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>things that you're supposed to practice</td>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>just the way it should be done</td>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exactly as you should be standing</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Group</td>
<td>Duration in cs.</td>
<td>Type of Group</td>
<td>Duration in cs.</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1-syllable</td>
<td></td>
<td>7-syllable</td>
<td></td>
</tr>
<tr>
<td>gét an'gt nnyhere</td>
<td>14</td>
<td>started heer at the univ-</td>
<td></td>
</tr>
<tr>
<td>hér wórking</td>
<td>22</td>
<td>vérsity</td>
<td>107</td>
</tr>
<tr>
<td>hé s'tarted</td>
<td>26</td>
<td>hâd wánted mé to bë a</td>
<td></td>
</tr>
<tr>
<td>nòw álso</td>
<td>26</td>
<td>missionây</td>
<td>116</td>
</tr>
<tr>
<td>2-syllable</td>
<td></td>
<td>dôn't fèel as if I cöuld gô</td>
<td>120</td>
</tr>
<tr>
<td>ány mán</td>
<td>30</td>
<td>I wánt to bë the oné up</td>
<td></td>
</tr>
<tr>
<td>'s a míñíster</td>
<td>37</td>
<td>thére</td>
<td>122</td>
</tr>
<tr>
<td>stúdent téáching</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gôné to Bâtíst</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-syllable</td>
<td></td>
<td>9-syllable</td>
<td></td>
</tr>
<tr>
<td>gô to a fôreign</td>
<td>46</td>
<td>décided thât we wëre go-ing to go into</td>
<td>142</td>
</tr>
<tr>
<td>héIsn't pérfect</td>
<td>52</td>
<td>állthing evëry happëns to</td>
<td></td>
</tr>
<tr>
<td>gët tröugh to thém</td>
<td>52</td>
<td>the húsband</td>
<td>144</td>
</tr>
<tr>
<td>sënior ríght nòw</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-syllable</td>
<td></td>
<td>5-syllable</td>
<td></td>
</tr>
<tr>
<td>forgët abôut the missionây</td>
<td>48</td>
<td>állnyone in my fâmîly</td>
<td>70</td>
</tr>
<tr>
<td>missionây fïeld</td>
<td>58</td>
<td>álso a pôssíbìlity</td>
<td>74</td>
</tr>
<tr>
<td>plán to go to séminary</td>
<td>74</td>
<td>állthing thât's worth-whîle</td>
<td>92</td>
</tr>
<tr>
<td>thât's all I cöuld dô</td>
<td>76</td>
<td>álsowent to my chûrch</td>
<td>92</td>
</tr>
<tr>
<td>5-syllable</td>
<td></td>
<td>6-syllable</td>
<td></td>
</tr>
<tr>
<td>àsnyone in my fâmîly</td>
<td>70</td>
<td>dîdn't think mâybe I cöuld</td>
<td>98</td>
</tr>
<tr>
<td>álso a pôssíbìlity</td>
<td>74</td>
<td>gët</td>
<td></td>
</tr>
<tr>
<td>állthing thât's worth-whîle</td>
<td>92</td>
<td>shë's accep'ted the idéà</td>
<td>100</td>
</tr>
<tr>
<td>álso went to my chûrch</td>
<td>92</td>
<td>thém so thëy cöuld under-</td>
<td></td>
</tr>
<tr>
<td>6-syllable</td>
<td></td>
<td>stánd</td>
<td>104</td>
</tr>
<tr>
<td>dîdn't think mâybe I cöuld</td>
<td></td>
<td>I wánt to do this wôrk álso</td>
<td>104</td>
</tr>
<tr>
<td>Type of Group</td>
<td>Duration in cs.</td>
<td>Type of Group</td>
<td>Duration in cs.</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1-syllable</td>
<td></td>
<td>7-syllable</td>
<td></td>
</tr>
<tr>
<td>nó spáre</td>
<td>27</td>
<td>cóok was a very very níce</td>
<td>75</td>
</tr>
<tr>
<td>spáre tíre</td>
<td>28</td>
<td>pèople as I said there were</td>
<td>fíve</td>
</tr>
<tr>
<td>óne súpermarket</td>
<td>30</td>
<td>famíliarize me with the contróls</td>
<td>110</td>
</tr>
<tr>
<td>áll náavigable</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-syllable</td>
<td></td>
<td>8-syllable</td>
<td></td>
</tr>
<tr>
<td>óut of tówn</td>
<td>38</td>
<td>fírst thing I noticed was it</td>
<td>had nó</td>
</tr>
<tr>
<td>tówn we stáyed</td>
<td>44</td>
<td>shòwed him that I really didn't táke</td>
<td>116</td>
</tr>
<tr>
<td>Kníghtsbridge Station</td>
<td>50</td>
<td>bús left at e leven twenty-twó</td>
<td>128</td>
</tr>
<tr>
<td>hígh-class bútcher</td>
<td>52</td>
<td>éach pub cated to a dif-férent sèction</td>
<td>134</td>
</tr>
<tr>
<td>3-syllable</td>
<td></td>
<td>9-syllable</td>
<td></td>
</tr>
<tr>
<td>státion to píck</td>
<td>44</td>
<td>Hítchen as I say was a</td>
<td>town of tén</td>
</tr>
<tr>
<td>pick up the cár</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unívér(s)i(y) of London</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Òxford and Cámbridge</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-syllable</td>
<td></td>
<td>10-syllable</td>
<td></td>
</tr>
<tr>
<td>órd(i)narily dó</td>
<td>50</td>
<td>provísioon mERCHANT to buy</td>
<td>a bottle of pòp</td>
</tr>
<tr>
<td>álter it was óver</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>áll very well dóné</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>give me a short tóur</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-syllable</td>
<td></td>
<td>11-syllable</td>
<td></td>
</tr>
<tr>
<td>ëveryone goes to béd</td>
<td>60</td>
<td>súpermarkets always smell heavily of sóap</td>
<td>162</td>
</tr>
<tr>
<td>Cámbridge and th(e)</td>
<td>Unívérsity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>twénty and the last bús</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cómes from having the</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>govérnment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-syllable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>óver we rénted a cár</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I want to go in a sòré</td>
<td>87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sérvice angle they don't cár</td>
<td>101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cábbie green that have been cóked</td>
<td>102</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Subject 5

<table>
<thead>
<tr>
<th>Type of Group</th>
<th>Duration in cs.</th>
<th>Type of Group</th>
<th>Duration in cs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-syllable</strong></td>
<td></td>
<td><strong>7-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>cute í</td>
<td>17</td>
<td>mariage was never on my mind</td>
<td>84</td>
</tr>
<tr>
<td>first girl</td>
<td>26</td>
<td>expectations of what it would be like</td>
<td>88</td>
</tr>
<tr>
<td>five six</td>
<td>34</td>
<td>probably end up with a major</td>
<td>92</td>
</tr>
<tr>
<td>had had</td>
<td>36</td>
<td>stáff announcer there last semester</td>
<td>96</td>
</tr>
<tr>
<td><strong>2-syllable</strong></td>
<td></td>
<td><strong>8-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>two years old</td>
<td>28</td>
<td>she's been working there about a year</td>
<td>81</td>
</tr>
<tr>
<td>each is speech</td>
<td>40</td>
<td>mother came over when she was eight</td>
<td>93</td>
</tr>
<tr>
<td>some time back</td>
<td>42</td>
<td>everything seems to be falling right</td>
<td>113</td>
</tr>
<tr>
<td>eight years old</td>
<td>43</td>
<td>exactly in what you would call a diligent</td>
<td>132</td>
</tr>
<tr>
<td><strong>3-syllable</strong></td>
<td></td>
<td><strong>9-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>sixth of October</td>
<td>41</td>
<td>planning on getting married in October</td>
<td>92</td>
</tr>
<tr>
<td>back in the thirties</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>turned against speech</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>speech altogether</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4-syllable</strong></td>
<td></td>
<td><strong>10-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>apparently it happened</td>
<td>36</td>
<td>think she's go(i)ng to do it in public address</td>
<td>104</td>
</tr>
<tr>
<td>stayed for about a month</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English is my minor</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>feel that I'm compatible</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5-syllable</strong></td>
<td></td>
<td><strong>11-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>I didn't enjoy that thought about turning anyway</td>
<td>61</td>
<td>teaching assignment at Jennings Junior High</td>
<td>151</td>
</tr>
<tr>
<td>integrated with English debate just about the best</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>6-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>everybody seems to ask categories in the speech</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I looked around quite a bit half as much as I did this</td>
<td>79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Subject 6

<table>
<thead>
<tr>
<th>Type of Group</th>
<th>Duration in cs.</th>
<th>Type of Group</th>
<th>Duration in cs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mísnómer</td>
<td>21</td>
<td><strong>7-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>sáme conflict</td>
<td>24</td>
<td>very valuable expérience</td>
<td>78</td>
</tr>
<tr>
<td>thát's ít</td>
<td>25</td>
<td>fáther and I don't get along</td>
<td>91</td>
</tr>
<tr>
<td>dó só</td>
<td>26</td>
<td>cónt(i)nuity director quit</td>
<td>98</td>
</tr>
<tr>
<td><strong>2-syllable</strong></td>
<td></td>
<td>jób in the Akron public school</td>
<td>105</td>
</tr>
<tr>
<td>thing thát's álways</td>
<td>27</td>
<td><strong>8-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>gó again</td>
<td>32</td>
<td>educación as there is anywhere else</td>
<td>86</td>
</tr>
<tr>
<td>sôme proficieny</td>
<td>34</td>
<td>educación is very rewarding</td>
<td>97</td>
</tr>
<tr>
<td>I look fórward</td>
<td>36</td>
<td>misnómer thát's going around the spéech</td>
<td>110</td>
</tr>
<tr>
<td><strong>3-syllable</strong></td>
<td></td>
<td>lót more money doing other thíngs</td>
<td>111</td>
</tr>
<tr>
<td>following pólítics</td>
<td>37</td>
<td><strong>9-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>sómething to púsh</td>
<td>43</td>
<td>responsibility for somebody else</td>
<td>94</td>
</tr>
<tr>
<td>áctive participation</td>
<td>51</td>
<td>thén I'd like to start working on my Mástér's</td>
<td>104</td>
</tr>
<tr>
<td>fífteen doors ópen</td>
<td>71</td>
<td>búsiness department has been about sîx</td>
<td>136</td>
</tr>
<tr>
<td><strong>4-syllable</strong></td>
<td></td>
<td><strong>10-syllable</strong></td>
<td></td>
</tr>
<tr>
<td>líke it very múch</td>
<td>51</td>
<td>éver been int(e)rested in</td>
<td>110</td>
</tr>
<tr>
<td>véry immatúre</td>
<td>48</td>
<td>radio wórkshop</td>
<td></td>
</tr>
<tr>
<td>álways gotten mé</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>proficieny in spéech</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hôpe I can get a jób</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>participation in pólítics</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I think thér'es as much pólítics</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hônestly say I sléep</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>6-syllable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pólítics in educátion</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nót that I don't wánt to téach</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>jób that I couldn't affórd</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>áverage this yár in the búsiness</td>
<td>97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
AUTOBIOGRAPHY

Domitila Domenech de Belaval was born to Puerto Rican and American parents in Ponce, Puerto Rico, in 1912. She was educated for the most part in the public schools of the island. She received a B. S. degree in English and French from Elmira College in 1933, an M. A. degree in comparative literature from Columbia University in 1947, and a Ph. D. degree in linguistics from Louisiana State University in 1966.

Since 1933 she has taught English at the University of Puerto Rico and is at present a member of the Department of English in the College of Humanities of that university.

She is married to Edgar S. Belaval, a lawyer, and lives at 612 Belaval Street, Santurce, Puerto Rico.
EXAMINATION AND THESIS REPORT

Candidate: Belaval, Domitila Domenech de

Major Field: Linguistics

Title of Thesis:

An Investigation of Isochronism in the Rhythm of American English Speech

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

May 2, 1966