DOCUMENT RESUME

ED 034 346 | 40 | EC 004 721

AUTHOR	Foulke, Emerson
TITLE	The Comprehension of Rapid Speech by the Blind: Part III. Final Report.
INSTITUTION	Louisville Univ., Ky. Non-Visual Perceptual Systems Lab.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-5-1050
PUB DATE	1 Sep 69
GRANT	OEG-4-10-127
NOTE	168p.
EDRS PRICE	EDRS Price MF-$0.75 HC-$8.50
DESCRIPTORS	Age Differences, Attitudes, *Comprehension, Electromechanical Aids, *Exceptional Child Research, Intelligence Differences, Listening Comprehension, Oral Reading, Reading Speed, *Speech Compression, *Visually Handicapped

ABSTRACT

Accounts of completed and ongoing research conducted from 1964 to 1968 are presented on the subject of accelerated speech as a substitute for the written word. Included are a review of the research on intelligibility and comprehension of accelerated speech, some methods for controlling the word rate of recorded speech, and a comparison of electromechanical and dichotic speech compression. Further discussions presented are the following: reaction time as an indicator of the intelligibility of time compressed words; the comprehension of compressed speech; listening comprehension as a function of intelligibility and word rate; a survey of the acceptability of rapid speech; preference of listening rate of college students of moderately difficult literary material; the influence of age, grade, and intelligence on comprehension of compressed speech; the rate of oral reading; the Louisville conference on time compressed speech; and summaries of other experiments. (JM)

BL 5-1050 PA-40

OE/BEH

ED034346

THE COMPREHENSION OF RAPID SPEECH BY THE BLIND: PART III

By

Emerson Foulke

Final Progress Report
Project No. 2430
Grant No. OE-4-10-127

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Covering Period: March 1, 1964 - June 30, 1968

A publication of the:

NON-VISUAL PERCEPTUAL SYSTEMS LABORATORY
GRADUATE SCHOOL
UNIVERSITY OF LOUISVILLE
LOUISVILLE, KENTUCKY 40208

September 1, 1969

EC 004 721E

ED034346

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FINAL

PROGRESS REPORT

Project No. 2430

Grant No. OE-4-10-127

THE COMPREHENSION OF RAPID SPEECH BY THE BLIND: PART III

Emerson Foulke

Non-Visual Perceptual Systems Laboratory
Graduate School
University of Louisville
Louisville, Kentucky 40208

Date submitted: September 1, 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FOREWARD

As human beings, we depend heavily, and perhaps more than we realize upon spoken language for the communication that supports daily living. Because aural communication is such an integral part of our lives, we have tended to take it for granted, and have not subjected it to the same kind of scrutiny that has been applied to recognized communication systems, such as the writing and reading of language. However, in recent years, the communication system, in which spoken language is the signal, has begun to receive more attention by educators and researchers. Consider, for instance, the courses now offered in many colleges and universities for the improvement of listening skills.

A special interest in communication by means of spoken language has been expressed by those who, for whatever reason, must place extraordinary reliance upon listening in order to communicate. Blind school children, for instance, depend to a considerable extent on listening to recorded spoken language because they do not have access to the communication system built around the print letter code, and because the rate at which braille is read is too slow to be practical in many situations.

One consequence of the increased interest in the process of aural communication has been a significant advance in the technology associated with the recording and reproduction of speech. As is true in the case of visual reading, a variable of obvious interest to those concerned with the process of aural communication, is the rate at which it occurs. Without special intervention, aural communication is governed by the rate at which speakers produce words. However, certain advantages might be gained if this rate could be altered. If it could be increased without a sacrifice in comprehension, the savings in time might be quite valuable to those who must depend upon aural communication. The ability to achieve selective reduction in the rate of communication might prove useful in educational settings such as foreign language classes, typing classes, remedial reading classes, etc.

The first method of altering the rate of recorded speech to receive the attention of investigators was the reproduction of a tape or record at a different speed than the speed used during recording. However, although this method achieves the desired effect as far as word rate is concerned, its inherent distortions seriously limit its usefulness. Fortunately, another method, pioneered by Dr. Grant Fairbanks (Fairbanks, Everitt, & Jaeger, 1954) at the University of Illinois, was introduced. This is a method in which, instead of reproducing an entire recording, periodic samples are reproduced and abutted in time. The duration of the samples that are not reproduced is brief enough so that the listener is not aware of their deletion. The result is speech that is reproduced in less than the original production time without distortion in vocal pitch or quality. With this method, the time required for the reproduction of a recording can be increased by repeating, rather than deleting periodic samples of the recording. The result is the same -- a change in word rate without distortion in vocal pitch or quality.

The ability to vary the time required for the reproduction of recorded speech without introducing serious distortion has stimulated a great deal of research concerning the effect on word intelligibility and listening comprehension of reproducing speech at some rate other than its natural rate of production. The results of many experiments support the conclusion that the word rate of recorded speech can be moderately increased without a significant loss in listening comprehension. Because of these findings, many people have begun to give serious consideration to a useful role for accelerated recorded speech in many educational settings. Programs organized around the needs of blind children constitute obvious examples.

The increased interest of those who wish to make practical use of the ability to control and vary speech rate has provided additional stimulation for researchers, with the result that there has been a rapid growth in the number of research projects exploring the educational significance of the ability to regulate speech rate. Since 1961, the Office of Education has supported a research project at the University of Louisville, a major objective of which has been the development of accelerated recorded speech, compressed in time by the sampling method, as a useful tool in the education of blind children. Research conducted in connection with this project has included investigations of the effect of the amount and method of time compression on the intelligibility of single words and the comprehension of connected discourse, the comprehension of connected discourse as a function of word rate with parameters such as difficulty of listening selection, age, sex, intelligence, and educational level of Ss, retention of the learning resulting from listening to accelerated speech, training experiences intended to promote better comprehension of accelerated speech, and the suitability of time compressed recorded speech for use in the aural reading of educational subject matter.

This volume contains accounts of research conducted during the support period extending from March 1, 1964, to June 30, 1968. Included are accounts of completed research, many of which have been reported elsewhere, and accounts of research in progress and preliminary investigations that have been suspended or discontinued for a variety of reasons.

ERIC

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF ILLUSTRATIONS

ERIC

CHAPTER I

A REVIEW OF RESEARCH ON THE INTELLIGIBILITY AND COMPREHENSION OF ACCELERATED SPEECH*

Emerson Foulke and

Thomas G. Sticht

Abstract

Time compressed or accelerated speech is speech which has been reproduced in less than the original production time. Such speech may prove to be useful in a variety of situations in which people must rely upon listening to obtain the information specified by language. It may also prove to be a useful tool in studying the temporal requirements of the listener as he processes spoken language. Methods for the generation of time compressed speech are reviewed. Methods for the assessment of the effect of compression on word intelligibility and listening comprehension are discussed. Experiments dealing with the effect of time compression upon word intelligibility and upon the comprehensibility of connected discourse, and experiments concerned with the influence of stimulus variables, such as signal distortion, and organismic variables, such as intelligence, are reviewed. The general finding that compression in time has a different effect upon the comprehensibility of connected discourse than upon word intelligibility is discussed, and a tentative explanation of this difference is offered.

---

*The article in this chapter also appears as an article in the Psychological Bulletin, 1969, 72, No. 1, 50-62.

Accelerated speech is speech in which the word rate has been increased. Increasing the word rate reduces communication time for a given message. Hence, accelerated speech is often referred to as time compressed, or simply compressed speech.

Since the announcement by Fairbanks (Fairbanks, Everitt, & Jaeger, 1954) of a practical means for the time compression of recorded speech, there has been an interest in its use to enable blind people to read by listening at a rate that compares favorably with the silent visual reading rate (Iverson, 1956; Foulke, Amster, Nolan, & Bixler, 1962). More recently, time compressed speech has been considered for use as an audio aid in general education (Orr & Friedman, 1964; Friedman, Orr, Freedle, & Norris, 1966) and as a research tool for studying the auditory perception of language (Foulke & Sticht, 1967).

This paper is concerned with the communication problems produced by the time compression of speech. Various techniques for the acceleration of speech are described, methods for its evaluation are reviewed, and characteristics of the listener that may affect his perception of time compressed speech are discussed.

Methods for the Acceleration of Speech

Speaking Rapidly

Within limits, word rate is under the control of the speaker, and this method has been used by several investigators (Calearo & Lazzaroni, 1957; deQuiros, 1964; Enc & Stolurow, 1960; Fergen, 1955; Goldstein, 1940; Harwood, 1955; Nelson, 1948). This method has the virtue of simplicity and requires no special equipment. However, it is limited by the fact that only a moderate increase in the rate of articulation of speech sounds is possible. When the speaker increases his word rate by talking faster, there are changes in vocal inflection and intensity, and in the relative duration of consonants, vowels, and pauses (Kozhevnikov & Chistovich, 1965). When word rate is increased by methods that alter the rate of reproduction of recorded speech, these changes do not take place. The significance of this fact, with respect to word intelligibility or listening comprehension, has not yet been determined.

The Speed Changing Method

The word rate of a recorded message may be changed, simply by reproducing it at a different tape or record speed than the one used during recording. If the playback speed is slower than the recording speed, word rate is decreased, and the speech is expanded in time. If playback speed is increased, word rate is increased, and the speech is compressed in time. When word rate is compressed in this manner, there is a shift in the frequencies that constitute the voice signal, which is proportional to the change in tape or record speed. If the speed is doubled, the component frequencies will be doubled, and vocal pitch will be raised one octave. Speech compressed by the speed changing method has been examined in several experiments (Fletcher, 1929, pp. 292-294; Foulke, 1966a; Garvey, 1953b; Klumpp & Webster, 1961; Kurtzrock, 1957; McLain, 1962).

The Sampling Method

In 1950, Miller and Licklider demonstrated the signal redundancy in spoken words, by deleting brief segments of the speech signal. This was accomplished by a switching arrangement that permitted a recorded speech signal to be turned off periodically during its reproduction. They found that as long as these interruptions occurred at a frequency of ten times per second, or more, the interrupted speech was easily understood. The intelligibility of monosyllabic words did not drop below 90% until 50% of the speech signal had been discarded. Thus, it appeared that a large portion of the speech signal could be discarded without a serious disruption of communication. Garvey (1953b), taking cognizance of these results, reasoned that if the samples of a speech signal remaining after periodic interruption could be abutted in time, the result should be time compressed, intelligible speech, without distortion in vocal pitch. To test this notion, he prepared a tape on which speech had been recorded by periodically cutting out short segments of tape, and by splicing the ends of the retained tape together again. Reproduction of this tape achieved the desired effect. Garvey's method was, of course, too cumbersome for any but research purposes. However, the success of the general approach having been shown, an efficient technique for accomplishing it was not long to follow.

In 1954, Fairbanks, et al., published a description of an electromechanical apparatus for the time compression or expansion of recorded speech, which embodies a principle adumbrated by Gabor (1946, 1947). The Fairbanks apparatus reproduces periodic samples of a recorded tape. The unreproduced samples are brief enough so that a discarded sample cannot contain an entire speech sound, and the retained samples are abutted in time. Under these conditions, every speech sound in the original recording is sampled, and the result is a time compressed reproduction without alteration in vocal pitch. Using this apparatus, speech can be expanded in time by periodically repeating samples of a recorded tape. A computer may also be used for the time compression or expansion of speech by the sampling method (Scott, 1965). Whereas speech compressors of the Fairbanks type sample periodically and unselectively, use of a computer permits a variety of sampling rules. For instance, a computer might be programmed to dispose of empty time intervals between words, and to sample the time intervals occupied by words differentially, discarding larger fractions of those speech sounds with higher signal redundancy. Though, because of its flexibility, the computer may provide the most satisfactory method for the time compression or expansion of speech, at present, computer time is too expensive to justify the employment of a computer in this capacity for any but research purposes.

The time compression of speech may be accomplished by shortening or eliminating the natural pauses occurring in speech (Miron & Brown, 1968; Diehl, White, & Burk, 1959). This may be done manually by removing blank segments of a recorded tape, or by means of a computer, and the remaining speech may be compressed or uncompressed.

The technique of speech synthesis offers another possibility for the compression of speech in time (Campanella, 1967). The harmonic compressor, a device for the time compression of speech based on research performed at Bell Laboratories, is now under construction at the American Foundation for the Blind.

Methods for the Evaluation of Accelerated Speech

Some Procedural Problems

There is no common practice in specifying the amount of compression to which a listening selection has been subjected. This lack of uniformity can result in confusion, especially when the results of different studies are compared (Bellamy, 1966). The amount of compression may be specified by the percentage of the original recording time that is saved by reproducing the message at a faster word rate. Thirty percent compression means that 30% of the production time has been saved. Conversely, the fraction of original production time remaining after compression may be specified.

Alternatively, specification may be in terms of the acceleration of the original word rate, tape speed, or record speed. An acceleration of 1.5 means that the word rate after compression is 1.5 times the word rate before compression. In comparing these indices, it must be remembered that the relationship between them is not linear. For instance, whereas an increase in acceleration from 1.1 to 1.2 corresponds to an increase in compression from 9 to 17%, an increase in acceleration from 1.9 to 2.0 corresponds to a change in compression from 47 to 50%.

A problem common to both indices is that they do not indicate directly the word rate of compressed speech. The final word rates of two listening selections, compressed or accelerated by the same amount, will depend upon the rates of speaking before compression. There is considerable variability in the published estimates of word rate. Part of this variability is undoubtedly due to the difference between spontaneous, conversational word rate, and the word rate of oral reading. Nichols and Stevens (1957) found a conversational speaking rate of 125 wpm, while Johnson, Darley, and Spriestersbach (1963, p. 220) found a median oral reading rate of 176.5 wpm, and Foulke (1967) found a mean oral reading rate of 174 wpm. The oral reading rate is the rate that is relevant to the process under discussion since, in most cases, the speech that is compressed is recorded oral reading. However, the usefulness of average oral reading rates is limited. The rate of oral reading depends upon the nature of the material being read, and this kind of variability can be reduced by reporting syllable rate, rather than word rate (Carroll, 1967). The oral reading rate also depends upon the style of the individual reader. It varies considerably from reader to reader, and from sample to sample of the production of a given reader (Foulke, 1967).

There are reasons for believing that speech rate is the dimension of which listeners are aware. Johnson, et al., (1963, pp. 202-203) have summarized research supporting the conclusion that perception of the rate of speaking corresponds to the oral reading rate. Hutton (1954) found a logarithmic growth in perceived word rate as measured word rate was increased linearly.

A variety of initial or uncompressed word rates has been used in studies of the effect of time compression on listening comprehension (Fairbanks, Guttman, & Miron, 1957c; Goldstein, 1940; Foulke, et al., 1962). These studies indicated that a rapid decline in comprehension commences beyond a word rate of approximately 275 wpm regardless of the compression which may have been required to achieve that word rate. Thus, it seems advisable to describe compressed speech not only in terms of the amount of compression, but also in terms of word rate.

For certain purposes, such as the measurement of intelligibility, single words are compressed, and it is, of course, meaningless to speak of the word rate of a single word. In these cases, specification must be made in terms of compression or acceleration ratio.

The Measurement of Intelligibility

The ability to repeat a word, phrase, or short sentence accurately, is often taken as an index of the intelligibility of time compressed speech. A procedure typical of this approach is one in which words are compressed in time by some amount and presented, one at a time, to a listener. The listener's task is to reproduce them orally, or in writing, and his score is the correctly identified fraction of those words. This procedure is sometimes referred to as an articulation test (Miller, 1954, p. 60).

Disjunctive reaction time (RT) may also be taken as an index of intelligibility (Foulke, 1965a). The underlying rationale, in this case, is that reduced discriminability means reduced intelligibility. It has been shown that if stimuli are made more similar, and hence less discriminable, choice RT is increased (Woodworth & Schlosberg, 1954, p. 33). The procedure for testing intelligibility, under this approach, is to acquaint S with a list of response words. The words are then presented to S, one at a time, in random order, for identification. Subject indicates his choice with a discriminative response, for instance, pressing an appropriate response key. He can then be scored for speed and accuracy of reaction. The experiment is performed using words that have been compressed in time by several amounts, and changes in RT and/or accuracy are regarded as indicative of changes in intelligibility. The RT method may be more sensitive than other methods, since a change in the amount of compression may produce a change in RT to words which are discriminated without error.

Calearo and Lazzaroni (1957) report the use of a method for testing intelligibility, familiar to those in clinical audiology, in order to detect the effects of compression. The minimum intensity required for words to be intelligible, is determined for words at several compressions. Threshold intelligibility is defined as that intensity at which some percent (usually 50) of a list of words is correctly identified. If the threshold for intelligibility changes as the amount of compression is changed, it is concluded that compression has affected intelligibility.

Tests of Comprehension

In this approach to the evaluation of the effects of compression, the listener first hears a listening selection, compressed in time by some amount, and is then tested for comprehension of that selection. Any kind of test may be used, but researchers have, in most cases, preferred objective tests of specifiable reliability.

Wood (1965) dealt with the problems inherent in assessing the listening comprehension of young children by determining their ability to follow brief, verbal instructions, compressed in time. Instructions consisted of imperative statements, such as "buzz like a bee".

Some tests of listening comprehension may detect differences not detected by others, but this increased sensitivity may have been purchased at the cost of a loss in reliability, or in ease of test administration and scoring. Bellamy (1966) used both a multiple-choice test and an interview technique to determine the listening comprehension of a group of blind Ss, and a comparable group of sighted Ss. She reports that the interview technique revealed a difference in favor of the blind Ss not detected by the multiple-choice test. Friedman, et al., (1966) used short answer and essay tests to assess the comprehension of accelerated speech, and found no discernable trend in performance as a function of practice in listening to such speech. On the other hand, a multiple-choice test revealed considerable improvement. They also found a lack of correlation between the results of short answer and essay tests.

The Intelligibility of Time Compressed Speech

Characteristics of the Signal

1. The method of compression. The intelligibility of time compressed words depends, in part, upon the method used for compression. When a recording is played back at a speed that is enough faster than the recording speed to result in the compressed reproduction of a list of words in two-thirds of their original production time, there is a loss in intelligibility of 40% or more (Fletcher, 1929; Garvey, 1953b; Klumpp & Webster, 1961; Kurtzrock, 1957). On the other hand, Garvey (1953b) found only a 10% loss in the intelligibility of a list of words, each of which was reproduced in 40% of its original production time by means of his manual sampling method, and a 50% loss in intelligibility for words reproduced in 25% of the original production times. Kurtzrock (1957), using the electromechanical sampling method of Fairbanks, obtained an intelligibility score of 50% for a group of words reproduced in 15% of their original production times. Using the same method and similar materials, Fairbanks and Kodman (1957) obtained an intelligibility score of 57% for a group of words reproduced in only 13% of their original production times.

Compression by either the sampling or the speed changing method increases the rate at which the discriminable elements of speech occur. However, whereas the overall spectrum, the location of formants within that spectrum, and vocal pitch are unaffected by the sampling method, they are altered by the speed changing method, and these alterations are probably responsible for the difference in intelligibility between the two methods (Nixon, Mabson, Trimboli, Endicott, and Welch, 1968; Nixon and Sommer, 1968).

2. Intelligibility and the sampling rule. The message to be compressed may be conceived as consisting of a succession of temporal segments, called sampling periods. When speech is compressed by the sampling method, compression is accomplished by discarding a fraction of each sampling period, and by abutting in time the remainders of sampling periods. It is the retained fraction of the sampling period that determines the amount of compression. If 10 milliseconds (msec.) of a 20 msec. sampling period or 30 msec. of a 60 msec. sampling period are retained, the result is the same -- 50% compression. For any given sampling period, changing the

ERIC

fraction of the sampling period that is retained changes the amount of compression.

When the sampling method is used, the effect that a given amount of compression will have upon the intelligibility of words depends upon the duration of the discarded portion of the sampling period, and hence upon the duration of the sampling period itself. The duration of the discarded portion of the sampling period must be short relative to the duration of the speech sounds to be sampled. If it is not, a speech sound may fall entirely within the discarded portion of a sampling period, in which case, it is not sampled at all. Garvey (1953b) used discard intervals of 40, 60, 80, and 100 msec., to compress spondaic words to 50% of their original durations. He obtained corresponding intelligibility scores of 95, 96, 95, and 86%. In a two factor experiment in which five discard intervals and eight compressions were represented, Fairbanks and Kodman (1957) also found a substantial loss in intelligibility when the duration of the discard interval exceeded 80 msec. This was true at all eight compressions.

Cramer (1965) reports that when Ss use earphones to listen to speech that has been compressed in time by the sampling method, delaying the signal to one earphone by 7.5 msec. improves intelligibility. This delay provides what Cramer has called "binaural redundancy". If, as Garvey (1953a) suggests, it is the briefness of highly compressed speech sounds that makes them unintelligible, binaural redundancy may restore some intelligibility by increasing the effective duration of speech sounds.

Scott (1965) reports a favorable result when Ss use one earphone to listen to the normally retained samples of compressed speech, and the other earphone to listen, at the same time, to the normally discarded samples of the same compressed speech. He refers to such speech as "dichotic speech".

3. The rate of occurrence of speech sounds. Garvey (1953b) compared the intelligibility of words compressed in time by the sampling method with the intelligibility, reported by Miller and Licklider (1950), of words that had been interrupted periodically. Garvey's words and Miller and Licklider's words were treated alike in that portions of sampling periods were discarded. However, the retained samples

of Garvey's words were abutted to produce time compressed speech, while the retained samples of Miller and Licklider's words were not abutted, and the resulting speech, though interrupted, was not compressed in time. There was no difference between the intelligibility of time compressed words and interrupted words when 50% of each word was discarded. However, when 62% of each word was discarded, interrupted words were 40% more intelligible than time compressed words. Since the two groups of words were alike with respect to the amount of speech information that had been discarded, the poorer intelligibility of the time compressed words, when 62% of the speech information was discarded, was probably due to the increased rate of occurrence of speech sounds. Garvey used spondaic words, whereas Miller and Licklider used monosyllabic words. Results obtained by Henry (1966) suggest that if Garvey had used monosyllabic words, or if Miller and Licklider had used spondaic words, the difference in favor of interrupted speech would have been even more pronounced.

4. Intelligibility and linguistic factors. Kurtzrock (1957) found that compression by the speed changing method degraded the intelligibility of vowel sounds more than consonantal sounds, and that compression by the sampling method degraded the intelligibility of consonantal sounds more than vowel sounds. Garvey's Ss (1953a) rated the vowel sounds in words that had been compressed in time by the sampling method higher in "goodness" than consonantal sounds. In a study in which the number of phonemes per word was varied from three to nine, Henry (1966) found that increasing the number of phonemes improved the intelligibility of words that had been compressed in time by the sampling method. In a similar vein, Klumpp and Webster (1961) found short phrases, compressed in time by the speed changing method, to be more intelligible than single words. The findings of Henry, and of Klumpp and Webster, are probably explained by the increased number of cues available to Ss because of the redundancy in polyphonemic words and short phrases, and could have been predicted from the finding of French and Steinberg (1947) that speech is understandable when composed of syllables that are only 67% intelligible.

Characteristics of the Listener

1. Intelligibility and prior experience. Fairbanks and Kodman (1957) found a group of words compressed by several amounts to be more intelligible than a similar group of words in which the same amounts

of speech information hac been discarded by interrupting them in the manner of Miller and Li klider. However, the Ss of Fairbanks and Kodman had received ~tensive familiarization with the words to be identified before the tests were made, whereas the Ss of Miller and Licklider were relatively naive.

Miller and Licklider (1950), using interrupted words, and Garvey (1953a), using words compressed in time by the sampling method, found that repeated exposure to such words improves their intelligibility.

If a group of listeners agree that a particular speech sound in a word that has been compressed in time by the sampling method is unrecognizable, it may fairly be concluded that the difficulty lies with the signal itself. However, Garvey found that Ss disagreed about the speech sounds that were rendered unintelligible by compression of the words in which they occurred. Garvey explained this finding in terms of the differential exposure of Ss to the words in question. In this connection, Henry (1966) found a positive relationship between word frequency in general language, as revealed in the Thorndike and Lorge (1944) word count and word intelligibility.

2. Intelligibility and hearing loss. There appear to be no differential effects of time compression upon the intelligibility scores of normal hearing Ss and patients having conductive or sensorineural hearing losses (Calearo & Lazzaroni, 1957; Bocca & Calearo, 1963; deQuiros, 1964; Luterman, Welsh, & Melrose, 1966; Sticht & Gray, in press). However, aged patients, some with diffuse cerebral pathology (Calearo & Lazzaroni, 1957; Sticht & Gray, in press), and patients with temporal lobe lesions (Bocca & Calearo, 1963; deQuiros, 1964) required greater intensity for threshold intelligibility and showed a higher error rate with supra-threshold words when compression was increased. The latter was true for aged Ss having normal hearing or sensorineural hearing losses (Sticht & Gray, in press). Apparently, the changes accompanying aging reduce the rate at which speech information can be processed.

Factors Affecting the Comprehension of Time Compressed Speech

Stimulus Variables

1. Comprehension and word rate. Within the range extending from 126 to 272 wpm, Diehl, et al., (1959) found listening comprehension

to be unaffected by changes in word rate. In the range bounded by 125 and 225 wpm, Nelson (1948) and Harwood (1955) found a slight but insignificant loss in listening comprehension as word rate was increased. Fairbanks, et al., (1957c) found little difference in the comprehension of listening selections presented at 141, 201, and 282 wpm. Thereafter, comprehension, as indicated by percent of test questions correctly answered, declined from 58% at 282 wpm to 26% at 470 wpm, a level of performance near chance. Foulke, et al., (1962), using both literary and technical listening selections, found listening comprehension to be only slightly affected by increasing word rate in the range bounded by 175 and 275 wpm. However, in the range extending from 275 to 375 wpm, they found an accelerating loss in listening comprehension as word rate was increased. Foulke and Sticht (1967) found a 6% loss in comprehension between 225 and 325 wpm, and a loss of 14% between 325 and 425 wpm. The three studies just cited are in agreement regarding the finding that as word rate is increased beyond a normal word rate, there is initially a moderate linear decline in comprehension, followed by an accelerating decline.

Simple comprehension scores do not take into account the learning time that is saved when speech is presented at an increased word rate. Such an allowance may be made by dividing the comprehension score by the time required to present the listening selection. This index of learning efficiency expresses the amount of learning per unit time. Using such an index, Fairbanks, et al., (1957c), Enc and Stolurow (1960), and Foulke, et al., (1962) found that learning efficiency increased as word rate was increased until a word rate of approximately 280 wpm was reached. In a similar approach, Enc and Stolurow (1960) computed an index of the efficiency of retention.

The word rate at which a listening selection is presented apparently has no special effect on the rate at which forgetting occurs. Enc and Stolurow (1960), Friedman, et al., (1966), and Foulke (1966b), performed studies in which tests of the comprehension of listening selections presented at several word rates were made after several retention intervals. In general, these studies support the conclusion that differences in the course of forgetting are due to differences in original learning. Of course, as has already been shown, the amount of original learning is, in part, a function of the word rate at which a listening selection is presented.

ERIC

2. Comprehension and the method of compression. McLain (1962) and Foulke (1962), using Ss who were naive with respect to compressed speech, and unaccustomed to reading by listening, compared the comprehension of a listening selection compressed by the sampling method to a rate of 275 wpm with the comprehension of the same selection compressed to the same word rate by the speed changing method. In both instances, a slight but statistically significant advantage was found for the sampling method. However, in a similar experiment in which blind children, who were accustomed to reading by listening, served as Ss, Foulke (1966a) found no statistically significant difference in favor of either method.

The finding that the obvious superiority of the sampling method, when the comparison is based upon a test of the intelligibility of single words, is not observed when the comparison is based upon a test of the comprehension of connected discourse, is of considerable interest. It suggests that some other factor, such as the rate at which words occur, is also involved in determining the comprehension of accelerated speech. A satisfactory explanation of such comprehension must, therefore, take into account the perceptual and cognitive processes of the listener.

3. Comprehension and the difficulty of the compressed material. The extent to which the comprehension of a listening selection is affected by compression in time may depend upon its difficulty. However, before this question can be examined satisfactorily, a method must be developed for determining the difficulty of a listening selection.

Using one normal and four accelerated word rates, Foulke, et al., (1962) measured the comprehension of a scientific selection and a literary selection. In each case, performance on a test containing multiple-choice items covering the listening section constituted the evidence for listening comprehension. Comprehension of the scientific selection was poorer than comprehension of the literary selection at a normal word rate, suggesting that it was relatively more difficult. As word rate was increased, comprehension of the scientific selection did not decline as rapidly as comprehension of the literary selection. Although this interaction was significant, it was probably due to the fact that since comprehension scores for the scientific selection were lower at a normal word rate, the range in which they could vary was relatively smaller. Furthermore, the apparent difference in difficulty of the two selections may have been due, at least in part, to differences in the tests of listening comprehension employed. Certainly, the apparent difficulty of a selection can be manipulated by the choice of items used in testing for its comprehension.

In an investigation of the effect of time compression on message units varying in difficulty, Fairbanks, et al., (1957c) distributed the 60 multiple-choice items of a test of listening comprehension equally among five categories of item difficulty. The listening selections covered by the test of comprehension were administered to several groups of Ss, each group experiencing a different accelerated word rate. Each S received five scores, determined by his responses to the items in each of the five test item categories. The mean score for each test item category decreased as the amount of compression in time was increased. They concluded that, assuming item difficulty to be a reflection of the difficulty of the message unit to which it pertained, the effect of time compression on listening comprehension, within the range explored, did not depend upon the difficulty of the listening material.

There are formulas for estimating what might be called the "absolute difficulty" of a selection. These formulas have generally been developed for material that is to be read visually (Dale & Chall, 1948; Flesch, 1948). However, it has often been assumed that the listening difficulty of a selection will be the same as its reading difficulty. The results of the experiment by Foulke, et al., (1962), suggest that this assumption may not be tenable. In this experiment, although comprehension test scores suggested that the scientific selection was relatively more difficult than the literary selection, they were estimated to be equal in difficulty by the Dale-Chall Formula for Readibility. Similar evidence is presented in a study reported by Enc and Stolurow (1960). They found considerable variability in the mean comprehension test scores of ten listening selections, presented at a normal word rate and a slightly accelerated word rate, in spite of the fact that the selections were rated as equal in difficulty by the Dale-Chall Formula. Of course, the formula may have failed to detect differences in listening difficulty because of a relatively large variance in the estimates of reading difficulty.

However, if the difficulty of an aurally received selection is not the same as the difficulty of that selection when visually received, the explanation may be that differences between the oral and the print display make it necessary for the reader to process them differently. The printed page is primarily a spatial display. It permits the kind of scanning that helps in understanding long, complex sentences. On the other hand, when information is specified by spoken language, it is displayed in a temporal dimension. The only sensory information available to the listener at any given instant is the information specified by the display at that instant. Unlike the visual reader, the listener must depend upon memory alone for the availability of speech that has already occurred. Furthermore, unlike the visual reader, he can exert no control over the order in which he encounters the syntactic and semantic components of sentences. The syntactical difference between two selections might be inconsequential when they are received visually, yet quite significant when they are received aurally. The formulas used for estimating reading difficulty (Dale-Chall, 1948; Flesch, 1948; Rodgers, 1962) are based on different considerations, and the estimates of difficulty yielded by these formulas may be expected to vary. However, there has been no comparative study of the extent to which the effect of word rate on listening comprehension depends upon the formula used to estimate difficulty. The finding of a systematic interaction between word rate and listening difficulty, as estimated by a particular formula, would seem to provide a kind of face validity for that formula.

4. Comprehension and the oral reader. Oral readers differ considerably with respect to vocal timbre, and of course, there are conspicuous sex differences in vocal pitch. Oral readers also differ with respect to such factors as average word rate, and variability in word rate, pitch, and loudness. Such factors combine to define the personal, oral reading style. In a preliminary experiment, Foulke (1964a) explored the extent to which oral reading style interacts with word rate in determining listening comprehension. Three renditions of a listening selection, each read by a different reader (two males and one female), were presented to three groups of college students at a normal word rate, and to three comparable groups at a word rate that was increased to 275 wpm by the sampling method. After exposure to the listening selection, all Ss took a test of listening comprehension. Significant differences in listening comprehension were associated with the reader variable, and with the word rate variable, but the reader's effect on listening comprehension did not depend upon the word rate at which the selection was presented.

Listener Variables That Affect Listening Comprehension

Foulke (1964a) has called attention to the considerable variation in the ability of listeners to comprehend accelerated speech. Several experiments have been reported in which there has been an effort to determine those characteristics of the listener that may contribute to the ability to comprehend accelerated speech.

1. The sex of the listener. Comparisons of male and female listeners have revealed no sex related differences in listening comprehension, for word rates ranging from 174 to 475 wpm (Foulke & Sticht, 1967; Orr & Friedman, 1964).

2. The listener's age and educational experience. Fergen (1955) and Wood (1965) found a positive relationship between the age-grade level of school children and their ability to comprehend accelerated speech. Together, their experiments included grades 1, 3, 4, 5, and 6.

3. The intelligence of the listener. In the case of children, the evidence presently available is not sufficient to permit a conclusion regarding the effect of intelligence on the comprehension of accelerated speech. Fergen (1955) found no relationship between the IQs of grade school children and their ability to comprehend accelerated listening selections. However, 230 wpm was the fastest word rate represented in her experiment. Wood (1965) found no relationship between the IQs of children in the primary grades and their ability to follow the instructions conveyed by short, imperative, time compressed statements. However, his procedures resemble more closely those used in testing for intelligibility. A more definite conclusion is possible in the case of adults. Fairbanks, *et al.*, (1957b, 1957c), Goldstein (1940), and Nelson (1948) have all found a positive relationship between intelligence and the ability to comprehend accelerated speech. The data of Fairbanks, *et al.*, (1957c) and Goldstein (1940) concur in showing a positive relationship between the intelligence of the listener and the magnitude of the decline in listening comprehension as word rate is increased. This relationship may be due, at least in part, to the fact that intelligent *S*s earn higher scores than less intelligent *S*s on comprehension tests of listening selections presented at normal word rates. Therefore, the scores they earn on tests of the comprehension of materials presented at accelerated word rates, have a larger range within which to vary.

4. The visual status of the listener. There are a priori grounds for expecting blind listeners to show better comprehension than sighted listeners. However, the research related to this question is meager and inconclusive. In an experiment performed by Hartlage (1963),

blind and sighted Ss did not differ with respect to their comprehension of listening selections presented at a normal word rate. Foulke (1964a) presented evidence that blind listeners comprehend time compressed listening selections better than sighted listeners.

5. Reading rate and listening rate. Those perceptual and cognitive processes that are responsible for individual differences in reading rate may also contribute to individual differences in the ability to comprehend accelerated speech. If this is true, fast readers should be able to comprehend speech at a faster word rate than slow readers. This hypothesis has been tested by Goldstein (1940), and by Orr, Friedman, and Williams (1965). In both experiments, a significant positive correlation was found between reading rate and the ability to comprehend accelerated speech. Of course, in all likelihood, a significant positive correlation would also have been found between reading rate and reading comprehension. In both experiments, it was also found that practice in listening to accelerated speech resulted in an improvement in reading rate.

Goldstein (1940), and Jester and Travers (1965) compared the comprehension resulting from listening to selections presented at several word rates with the comprehension resulting from reading the same selections at the same word rates. In both cases, comprehension declined as word rate was increased. Listening comprehension was superior to reading comprehension up to approximately 200 wpm, but inferior to reading comprehension thereafter. Simultaneous reading and listening at 350 wpm resulted in better comprehension than could be demonstrated with either mode of presentation alone.

6. Improving the comprehension of time compressed speech. In an experiment performed by Fairbanks, et al., (1957b), a mean comprehension score of 63.8% was obtained by Ss who listened to a selection presented at an uncompressed word rate at 141 wpm. Subjects who listened to the same selection, compressed by 50% to a word rate of 282 wpm, earned a mean comprehension score of 58%. A third group of Ss, who listened to two consecutive reproductions of the listening selection at 282 wpm, earned a mean comprehension score of 65.4%, which was slightly, but probably not significantly higher than the mean comprehension score resulting from a single exposure to the uncompressed selection. In a second study, by the same investigators (1957a), augmentations were written for selected facts in a listening selection.

The recorded version of the augmented selection was then compressed enough by the sampling method to produce a playback time equal to the playback time of the uncompressed and unaugmented selection. The objective was to determine whether or not comprehension could be improved by trading the temporal redundancy in the uncompressed version for the verbal redundancy in the augmented version. Analysis of the results revealed better comprehension only for the augmented sections of the listening selection. There was a decline in comprehension of the unaugmented sections. The explanation of this finding may be that Ss associated verbal redundancy with importance, and distributed their attention accordingly.

Several investigators have explored the possibility of improving the comprehension of accelerated speech by training. The simplest, and least sophisticated training experience that has been evaluated, is mere exposure. Voor and Miller (1965) exposed a group of Ss to five listening selections, presented at 380 wpm. Total listening time was 17.5 minutes. At the end of each selection, Ss were tested for listening comprehension. Mean comprehension scores increased from the first to the third selection, but did not change significantly thereafter. These results probably reflect a simple adjustment to the initially unfamiliar task of listening to accelerated speech.

Orr, Friedman, and Williams (1965) found a 29.3% increase in the comprehension of materials presented at 475 wpm, following several weeks of training in which Ss listened to selections, the word rates of which were increased in steps of 25 wpm from 325 to 475 wpm. However, since there was no control group that received training in listening for comprehension at a normal word rate, it is not possible to attribute their results unequivocally to practice in listening to accelerated speech. The improvement may have been due simply to practice in listening for comprehension.

In this regard, Foulke (1964a), using blind Ss who can safely be presumed to have had years of experience in listening for comprehension, measured their comprehension of speech presented at 350 wpm, before and after training. Training consisted of approximately 25 hours of exposure to (a) speech at a constant rate of 350 wpm, (b) speech that was gradually increased from a normal word rate to a final word rate of 350 wpm, (c) the same as (a) but with frequent pauses for questioning about the material just heard, and, (d) the same as (b) but with frequent pauses for questioning about material just heard. There were no significant differences between pre- and post-training test scores for any of the treatment groups.

Friedman, et al., (1966) compared the comprehension test scores of Ss given 35 hours of massed practice in listening to accelerated speech with the comprehension test scores of Ss who received from 12 to 14 hours of distributed practice in listening to accelerated speech. They concluded that the comprehension demonstrated by the distributed practice group was as good as, or better than, the comprehension demonstrated by the massed practice group.

From the research reviewed above, it is clear that an adequate training experience for improving the comprehension of accelerated speech has yet to be found. Simple exposure, at least in the amounts so far tested, is not adequate.

Conclusion

It is possible to provide a fairly accurate description of the relationship between word rate and listening comprehension on the basis of the experimental results that have been reviewed. There are two general classes of results which, when taken together, suggest that the relationship between word rate and listening comprehension is structured by more than one underlying process. First, there are those studies in which listening comprehension has been measured at various word rates (see Stimulus Variables, pg. 11). When these studies are considered collectively, the relationship that emerges is one in which listening comprehension declines at a slow rate as word rate is increased, until a rate of approximately 275 wpm is reached, and at a faster rate thereafter.

In the second class of studies, intelligibility has been determined for words compressed by various amounts (see Characteristics of the Signal, pg. 8). These studies are in general agreement regarding the finding that, when compression is accomplished by the sampling method, word intelligibility is not seriously degraded until a relatively large amount of signal information has been discarded. The finding that increasing the amount of compression has a different effect upon listening comprehension than upon word intelligibility suggests that decreased intelligibility is not, in itself, an adequate explanation for the loss in comprehension that is observed at faster word rates. One might expect decreased intelligibility to interfere with comprehension to some extent.

However, the listener's uncertainty regarding imminent speech is reduced because of his ability to estimate the sequential dependencies in meaningfully connected words and syllables, and there is a further substantial reduction in uncertainty when he has heard enough of a message to form a valid hypothesis about its contents. The reduction in message uncertainty should significantly counteract losses in word intelligibility, and the finding by French and Steinberg (1947), that listeners can understand messages composed with words whose syllables are only 67% intelligible, suggests that this is the case.

The increase in the rate at which comprehension declines beyond 275 wpm, suggests that when a certain critical word rate is reached, a factor in addition to signal degradation begins to determine the loss in comprehension. The understanding of spoken language implies the continuous registration, encoding and storage of speech information, and these operations require time. When the word rate is too high, words cannot be processed as fast as they are received, with the result that some speech information is lost. To put it another way, when channel capacity is exceeded, some of the input cannot be recovered at the output (Miller, 1953; 1956).

The explanation just suggested is, of course, tentative. A good deal of research on sentence, word, and syllable rate, and upon the amount of distribution of processing time in connected discourse, will be required in order to provide a more substantial basis for the hypothesis.

CHAPTER II

METHODS FOR CONTROLLING THE WORD RATE OF RECORDED SPEECH

by

Emerson Foulke

Abstract

Six methods for increasing speech rate are presented. They are as follows. 1. Speech at a rate that is faster than normal may be obtained by pacing an oral reader at a rate that is faster than his normal reading rate. 2. The word rate of recorded speech may be increased by reproducing a tape or record at a speed that is faster than the speed used during recording. 3. The word rate of recorded speech may be increased by an electromechanical device that reproduces consecutive samples of a recorded tape. 4. Consecutive sampling may also be accomplished by a computer. 5. The word rate of synthesized speech may be manipulated by instructions in the program followed by a speech synthesizer. 6. The harmonic compressor increases word rate by a method of frequency division without temporal alteration, and frequency restoration with temporal alteration.

There are several methods for increasing the word rate of recorded speech. None of these methods are completely free from distortion, and each method imposes its own, characteristic distortion. By now, a good deal of research has been accomplished in which one or more methods have been evaluated with respect to their effect on word intelligibility and/or listening comprehension. Though a review of such research is not within the scope of this article, summary statements of research

ERIC

findings will be made where appropriate, and pertinent references will be cited.

Before turning to the description of the various methods, a few remarks are in order regarding confusion in the terminology used in talking about recorded speech, the word rate of which has been increased. Any recorded speech that is reproduced in less time than the time required for its original production can be regarded as having been compressed in time. Hence, such speech is often called time compressed speech, or simply compressed speech. Since reproducing recorded speech in less time than the time required for its original production results in an increase in word rate, it is often called accelerated speech. Such speech has also been described as rapid speech or speeded speech. There has been an attempt on the part of some writers to employ these terms selectively in describing the products of the various methods. However, there has been no general agreement about which term should be used for the product of which method. In the present article, there is no need for such terminological differentiation, since the discussion will be primarily of the methods themselves, and not of their products. An attempt to secure agreement among researchers regarding the appropriate term for the product of each of the several methods might be a useful undertaking. In the absence of such agreement, it will continue to be necessary for writers to avoid referring to recorded speech, the word rate of which has been increased, without specifying the method by which this has been accomplished.

Speaking Rapidly

Increasing word rate by speaking rapidly is the only method presented in this paper that does not operate upon recorded speech. Its discussion is included here for the sake of completeness, and because the comparison of this method with other methods exhibits a class of variables that may have to be taken into account in producing comprehensible speech at an increased word rate.

Within limits, word rate is under the control of the speaker (Calearo & Lazzaroni, 1957; deQuiros, 1964; Enc & Stolurow, 1960; Fergen, 1955; Goldstein, 1940; Harwood, 1955; Nelson, 1948). This method requires no exotic apparatus. However, if the increased word rate that results from speaking rapidly is to be well controlled, the speaker must be trained, and he must be provided with feedback to regulate his speaking rate. This method has a distinct disadvantage. When a speaker attempts to operate his speech machinery at a rate that is much faster than normal, it begins to malfunction. That is, when the muscles involved in the articulation of speech sounds are made to respond too rapidly, the coordination of their action begins to deteriorate, with resulting errors in articulation. Furthermore, even below this critical limit, it is doubtful that a speaker can maintain a speaking rate that is faster than his normal rate for very long at a time.

As a speaker produces connected speech, he varies vocal pitch, vocal intensity, and the amount and distribution of pause time. Although there is, at present, an insufficient amount of research regarding the contribution of these variables to the comprehensibility of spoken language, it is a fair hypothesis that, in addition to the information contained in the words the speaker uses and in the order in which he arranges them, he specifies something about his message by the way in which he jointly manages pitch, intensity, and pause time. Goldman-Eisler (1956), for instance, has introduced the concept of cognitive rhythm, which she believes to be an essential feature of spoken language, and which is the result of the way in which a speaker distributes pause time in his speech production.

When a speaker attempts to speak more rapidly, there are departures from his characteristic use of pitch, intensity, and pause time (Goldman-Eisler, 1956). The sampling method (see pg. 24, ln. 15) preserves both pitch and intensity, and although it reduces the absolute amount of pause time, it preserves the apportionment of pause time in a speech production. The speed changing method (see pg. 23, ln. 35) like the sampling method, preserves vocal intensity and the apportionment of pause time. It elevates overall pitch, but preserves the relationship among the frequencies in the voice signal. What is preserved, and what is not preserved as speech is compressed, may prove to be an important consideration in evaluating the various methods of compression.

The Speed Changing Method

The word rate of recorded speech may be changed simply by reproducing a tape or record at a different speed than the one used during recording. If the playback speed is slower than the recording speed, word rate is decreased and the speech is expanded in time. If the playback speed is increased, the word rate is increased, and the speech is compressed in time. When speech is accelerated in this manner, there is a change in the frequencies that constitute the voice signal. This change is proportional to the change in tape or record speed. If playback speed is doubled, the component frequencies will be doubled, and vocal pitch will be raised one octave. Speech compressed by the speed changing method has been examined in several experiments (Fletcher, 1929, pp. 292-294; Foulke, 1966a; Garvey, 1953b; Klumpp & Webster, 1961; McLain, 1962). These experiments indicate that both the intelligibility of single words and the comprehension of connected discourse withstand only moderate compression in time before losses set in.

The Sampling Method

In 1950, Miller and Licklider demonstrated the signal redundancy in spoken words by deleting brief segments of the speech signal. This was accomplished by a switching arrangement which permitted a recorded speech signal to be turned off periodically during its reproduction. They found that as long as these interruptions occurred at a frequency of ten times per second or more, the interrupted speech was easily understood. The intelligibility of monosyllabic words did not drop below 90% until 50% of the speech signal had been discarded. Thus, it appeared that a large portion of the speech signal could be discarded without a serious disruption of communication.

Garvey (1953b) taking cognizance of these results, reasoned that if the samples of a speech signal remaining after periodic interruption could be abutted in time, the result should be time compressed intelligible speech without distortion in vocal pitch. To test this notion, he prepared a tape on which speech had been recorded by periodically cutting out short segments of tape and by splicing the ends of the retained segments of tape together again. Reproduction of this tape achieved the desired effect. Garvey's method was, of course, too cumbersome for any but research purposes. However, the success of the general approach having been shown, an efficient technique for accomplishing it was not long to follow.

In 1954, Fairbanks, et al., published a description of an electro-mechanical apparatus for the time compression or expansion of recorded speech, which embodies a principle adumbrated by Gabor (1946, 1947). In the Fairbanks apparatus, a continuous tape loop passes over a record head, used to place on this storage loop the signal that is to be compressed. Next, the tape passes over the sampling wheel, which reproduces samples of the signal that has just been recorded. Finally, it passes over an erase head that removes the signal from the storage loop so that it can be re-recorded on the next cycle. The sampling wheel is a cylinder, with four playback heads embedded in it, flush with its curved surface, and equally spaced around the curved surface. The tape, in passing over the curved surface of the sampling wheel, makes contact with approximately one-quarter of its surface. When the sampling wheel is stationary, and one of its heads is contacted by the moving tape, the signal on the tape is reproduced as recorded. However, when the apparatus is adjusted for some amount of compression, the sampling wheel begins to rotate in the direction of tape motion. Under these conditions, each of the four heads, in turn, makes and then loses contact with the tape. Each head reproduces the signal on the portion of the tape with which it makes contact. When, as it rotates, the sampling wheel has arrived at a position at which one head is just losing contact with the tape, while the preceding head is just making contact, the segment of tape that is wrapped around the sampling wheel between these two heads never makes contact with a reproducing head, and is therefore not reproduced. The segment of tape that is eliminated from the reproduction in this manner is always the same length, one-quarter of the circumference of the sampling wheel. The amount of speech compression depends upon the frequency with which these tape segments are eliminated, and this frequency depends, in turn, upon the rotational speed of the sampling wheel. The temporal value of the segments of tape that are not reproduced depends upon the speed of the storage loop, since this determines the amount of tape that will pass over a tape head during a given time interval. Since the sampling wheel rotates in the direction of tape motion, the speed of the storage loop, relative to the surface of the sampling wheel, is reduced, with the result that the frequencies in the retained samples of the original signal are lowered. The output of the compressor is recorded on tape, and this tape is reproduced at a speed that is enough faster than the recording speed to restore the lowered frequencies to their original values. The increase in the playback speed of this tape results in its reproduction in less than the original production time, and the result is time compressed speech that is not altered with respect to vocal pitch. In an alternate mode of operation, the tape or record player which supplies the signal to the record head that transfers it to the compressor's storage loop, may be speeded up enough to produce an elevation in the frequencies constituting the signal that is exactly compensated for by the lowering of frequencies which takes place during the sampling process. In this case, the output signal of the compressor is compressed in time without frequency distortion.

Speech may be expanded in time by reversing this process. The sampling wheel is rotated in a direction opposite to that of the storage loop, so that samples of the signal recorded on it are periodically repeated.

The speech compressor now manufactured by Mr. Wayne Graham* is based upon the Fairbanks design. Like the Fairbanks compressor, it makes use of a storage loop. The temporal value of the samples that are discarded during compression can be varied by changing speed of the storage loop. Operation of the Graham compressor requires two tape recorders -- one to provide its input, and one to receive its output. One of these recorders must be continuously variable in speed.

Mr. Anton Springer, relying upon the same basic principle, developed a compressor with a modified mode of operation**. In the Springer approach, the storage loop, the record head, and the erase head have been eliminated. Previously recorded tape passes from a supply reel over the surface of the sampling wheel to a take up reel. The tape is sampled in the manner just described. However, as the sampling wheel rotates in the direction of tape motion, the speed of the tape is increased by an amount sufficient to hold tape speed constant in relation to the surface of the sampling wheel over which it passes. Thus,

---

*Mr. Wayne Graham, Discerned Sound, 4459 Kraft Avenue, North Hollywood, California 91602.

**The current version of the Springer device, known as the Information Rate Changer, is distributed in this country by Infotronic Systems, Inc., 2 West 46th Street, New York, New York 10036.

the output of the Springer device is compressed in time, without distortion in vocal pitch. The temporal value of the samples discarded during compression by the Springer device is determined by the distance, along the curved surface of the sampling wheel, separating adjacent playback heads, and is not variable. Operation of a compressor of the Springer type requires a tape recorder to receive its output. In addition, another tape recorder is required to provide the tape transport function, since the commercially available compressors based on the Springer approach have not incorporated provisions for handling tape.

A computer may also be used for compressing speech by the sampling method (Scott, 1965). In this approach, speech that has been transduced to electrical form, for example, the output of a microphone or tape reproducing head, is temporally segmented by an analog-to-digital converter, and these segments are stored in the computer. The computer samples these segments according to a sampling rule for which it has been programmed; for example, discard every third segment. The durations of both retained and discarded samples can be varied over a wide range. The retained samples are abutted in time, and fed to the input of a digital-to-analog converter, and the signal at the output of this converter, compressed in time, is appropriate for transduction to acoustical form again.

Electromechanical compressors of the Fairbanks or Springer type are unselective with respect to the portions of a recorded signal that are discarded. Portions are discarded on a periodic basis, and may be deleted anywhere within or between words. It is quite unlikely that a given signal would be sampled in exactly the same way on two consecutive passes through such a device. With the computer, it is feasible to employ a variety of sampling rules. For instance, a computer might be programmed to dispose of empty time intervals between words, and to sample the time intervals occupied by words differentially, discarding larger fractions of those speech sounds with higher signal redundancy. From what has just been said, it would appear that the computer, because of its greater flexibility, offers the most satisfactory approach for the time compression of speech. This may ultimately prove to be the case. However, at present, computer time is too expensive to justify the employment of a computer in this capacity for any but research purposes.

Furthermore, although researchers such as Scott and Cramer* are working on the problem of writing programs for the differential samplings of speech signals, satisfactory programs have not yet been written.

Speech compressed by the sampling method has been evaluated with respect to word intelligibility (Fairbanks & Kodman, 1957; Foulke & Sticht, 1967; Garvey, 1953b; Kurtzrock, 1957) and listening comprehension (Fairbanks, et al., 1957c; Foulke, et al., 1962; Reid, 1968). In general, results have shown that whereas word intelligibility is relatively resistive to the effects of compression by the sampling method, listening comprehension begins to decline after moderate compression. Several investigators have tested training experiences intended to improve the comprehension of time compressed connected discourse (Foulke, 1964a; Orr, et al., 1965). Although the successful training experience has not yet been devised, Orr, et al., have reported encouraging results.

Other Methods for the Time

Compression of Speech

The technique of speech synthesis suggests another possibility for the production of accelerated speech without distortion in vocal pitch (Campanella, 1967). The speech synthesizer generates electrical analogs of the acoustical materials needed for the construction of speech sounds. A program of rules is provided for generating these analogs for the proper durations, at the proper intensities, and in proper conjunction or sequence. These rules may be varied to produce speech at any described rate. Though this method has, as yet, received little development, it should share with the computer the ability to shorten speech sounds in accordance with their signal redundancy.

---

*Dr. Robert Scott, 8604 Bunnell Drive, Potomac, Maryland 20854; Dr. H. Leslie Cramer, 156 Line Street, Cambridge, Massachusetts 02139.

Another device for the time compression of speech, now under development at the American Foundation for the Blind, is the harmonic compressor, an outgrowth of research conducted at the Bell Laboratories. In this approach, a speech signal is passed through an elaborate filtering network which divides the speech spectrum into a large number of narrow frequency bands. The portion of the signal appearing in each of these bands is then reduced in frequency by one-half, by means of multivibrator circuitry. The resulting signals are then combined again to produce speech, the frequencies of which have been reduced by one-half. If a recording of this speech is reproduced at twice the recording speech, the result is speech that has been compressed to 50% of the original production time, without a change in vocal pitch. Since the prototype of this compressor has only just been completed, there has been no opportunity to evaluate its output. A serious limitation of the harmonic compressor is that it cannot be adjusted for any desired amount of compression. It can only reduce the time required for the reproduction of a message by one-half.

CHAPTER III

A COMPARISON OF "DICHOTIC" SPEECH AND SPEECH COMPRESSED BY THE ELECTROMECHANICAL SAMPLING METHOD*

by

Emerson Foulke and

E. McLean Wirth

Abstract

An experiment was performed to compare the Fairbanks method of electromechanical speech compression and the computer sampling method resulting in dichotic speech, described by Scott, with respect to their effects on the intelligibility of phonetically balanced spoken words. Comparisons were made at five compressions in time: 47%, 44%, 41%, 39%, and 37% of original production time. The number of errors made in identifying words increased as the amount of compression was increased, but no significant difference in errors was associated with the method of compression used.

Recorded speech may be compressed in time by reproducing a succession of periodic, time abutted samples of the original recording. If the durations of the samples eliminated from such a reproduction are brief enough so that no critical feature of a speech signal can, by accident of sampling, fall entirely within a discarded sample, the result is time compressed, intelligible speech that is not altered with respect to vocal pitch or quality.

---

*The research described in this report was also reported by the junior author in her senior thesis, submitted to the Webster College, St. Louis, Missouri, 1968.

Such sampling may be accomplished manually (Garvey, 1953b), by cutting a recorded tape into segments, discarding some of the segments, and splicing the remaining segments together again. It may be accomplished more conveniently by a tape reproducer of the type described by Fairbanks, et al., (1954). Devices of the Fairbanks type reproduce periodic, time abutted samples of a recorded tape and, as before, the result is time compressed, intelligible speech, without distortion in vocal pitch or quality. (For a more complete description of this process, see pg. 25, ln. 4.)

A computer may also be used for the time compression of speech (Cramer, 1968; Scott, 1965). In this approach, the recorded speech signal is temporally segmented, some of the time segments are discarded according to a sampling rule for which the computer has been programmed, and the remaining segments, abutted in time, are reproduced as time compressed speech. (For a more complete description of this process, see pg. 27, ln. 10.)

In a scheme proposed by Scott (1967), the signal resulting from the process just described is applied to one earphone of a headset. The samples that would have been discarded in the kind of compressed speech described heretofore, are retained, abutted in time, and supplied to the other earphone. With this approach, for compressions in time of 50% or less, all of the recorded signal is preserved in the compressed reproduction. It is only rearranged temporally. For compressions greater than 50%, some of the signal must be discarded, but much more is preserved than when only one succession of samples is reproduced. Scott calls the product of this process "dichotic speech".

When speech is compressed by an electromechanical compressor of the Fairbanks or Springer type, a single file of time abutted samples is reproduced and this method will be referred to hereafter as the single file sampling method. When a computer is used to produce dichotic speech, two parallel files of time abutted samples are reproduced, and this method will be referred to hereafter as the double file sampling method.

When speech is compressed in time by discarding samples of the original signal, as the length of samples is reduced, the probability is reduced that a critical feature of a speech signal will fall entirely within a discarded sample (Garvey, 1953b). In designing a speech compressor, the physical parameters of the system must be adjusted to produce discard samples, the durations of which are short enough so that the probability of discarding a critical feature of a speech signal can safely be ignored. Two types of speech compressors have been developed for commercial distribution. One is based directly upon the Fairbanks scheme (for the Graham compressor, see footnote *, pg. 26, ln. 31). The other, based directly upon the Springer scheme, is the Information Rate Changer (see footnote **, pg. 26, ln. 33). The Fairbanks scheme permits adjustment of the duration of discarded samples. In the Springer scheme, this capability is sacrificed in the interest of convenience of operation*. In either case, however, samples are discarded, and there is some probability that one or more of these samples may contain a critical feature of a speech signal. Since the process resulting in dichotic speech discards none of the speech signal in the range of compression bounded by zero and 50%, the probability of discarding a critical feature of a speech signal should be reduced to zero. Consequently, a reasonable conjecture would be that, in the long run, words compressed by the process resulting in dichotic speech should be somewhat more intelligible than words compressed by discarding samples of the speech signal. The superior intelligibility of dichotic speech might not be manifested on any given comparison of the two alternative reproductions of a single word. However, as the length of the list of words used for such a comparison was increased, there would be an increased opportunity for the sampling accidents that can occur with the single file sampling method, and the relative superiority of dichotic speech should begin to emerge. Accordingly, an experiment was performed in which a list of words, compressed by the two methods just described, were compared with respect to intelligibility.

Method

Subjects

Sixty Ss, of both sexes, enrolled in introductory psychology classes at the University of Louisville, served in the experiment. Subjects had no obvious hearing defects, and little or no prior experience in listening to time compressed speech.

---

*The duration of the discarded samples produced by the Information Rate Changer, a currently available commercial device embodying the Springer scheme, is 30 msec.

Apparatus and Materials

A list of 100, phonetically balanced words was read orally by a professional reader in the Talking Book Studios of the American Printing House for the Blind, and recorded on magnetic tape by means of an Ampex tape recorder, model 300. This "master tape" supplied the input to a speech compressor of the Springer type, constructed at the University of Louisville, and to the computer used in preparing dichotic speech*. Since the samples discarded by the electromechanical speech compressor were 40 msec. in duration, the computer was adjusted so that the samples normally discarded, but retained by the computer for dichotic presentation, were 40 msec. in duration, too. The master tape was reproduced, by both methods, in 47%, 44%, 41%, 39%, and 37% of the original production time. If a recording of connected speech, occurring at the average oral reading rate of 175 wpm (see pg. 106, ln. 2), were subjected to these compressions, the resulting word rates would be 375, 400, 425, 450, and 475 wpm. Compressions in this range were chosen because earlier research (Garvey, 1953b; Fairbanks & Kodman, 1957; Kurtzrock, 1957) indicated that words presented at more moderate compressions would have been completely intelligible, with either kind of compression. The compressed reproductions were copied on magnetic tape for presentation in the experiment. In the case of dichotic presentation, the normally retained samples of the compressed signal were recorded on one track of a two-track stereo tape, while the normally discarded samples were recorded on the other track. Of course, only one track was required for recording the output of the electromechanical compressor. These tapes were reproduced, during the experiment, on a Revox tape recorder, model G36-III. The tape recorder was connected through a Pilot stereo preamplifier model 216A, and a Pilot stereo amplifier model SA-260 to a pair of Western Electric headphones, type ANB-H-1, equipped with ear cushions, and wired for stereophonic listening. When the tape containing speech compressed by the double file sampling method was reproduced, the file of samples recorded on one track of the tape was presented to one ear, and the file of samples

*Dichotic speech was prepared for this experiment at the National Security Agency, Fort George G. Meade, Maryland, by John Boehn, using methods developed by Dr. Robert Scott. Dr. Scott's assistance in arranging for the preparation of this material is sincerely appreciated.

recorded on the other track was presented to the other ear. When the tape containing words compressed by the single file sampling method was reproduced, the same signal was presented to both ears. The E monitored the experiment by listening to another pair of earphones, connected to an auxillary output on the tape recorder.

Procedure

The 60 Ss were divided into five groups, with 12 Ss in each group. Each group was tested with words presented at only one of the five compressions represented in the experiment. Six members of each group heard the first 50 words in the list, compressed by the double file sampling method. The remaining 50 words were compressed by the single file sampling method. For the other six members in each group, the first 50 words in the list were compressed by the single file sampling method, while the remaining words were compressed by the double file sampling method and presented as dichotic speech. This precaution was taken to control for the possibility that some words may have been treated more favorably by one method or the other. To control for the possibility of an effect due to order, three of the Ss in each sub-group heard words compressed by the double file sampling method, followed by words compressed by the single file sampling method. The order of presentation was reversed for the remaining three Ss in each sub-group.

Subjects were tested one at a time. Each S wrote the words he thought he heard on an answer sheet in numbered answer spaces. Approximately five seconds elapsed between the onsets of consecutive words. Subjects were instructed to guess if they were uncertain about a word.

Results

At each fraction of original production time represented in the experiment, two scores were determined for each S -- the number of words compressed by double file sampling that were missed, and the number of words compressed by single file sampling that were missed. Means and standard deviations of error scores are shown in Table 3.1. The influence of the method of compression upon the relationship between the amount of compression and error frequency is graphed in Figure 3.1. In this figure, the fraction of original production time required for compressed reproduction, at each

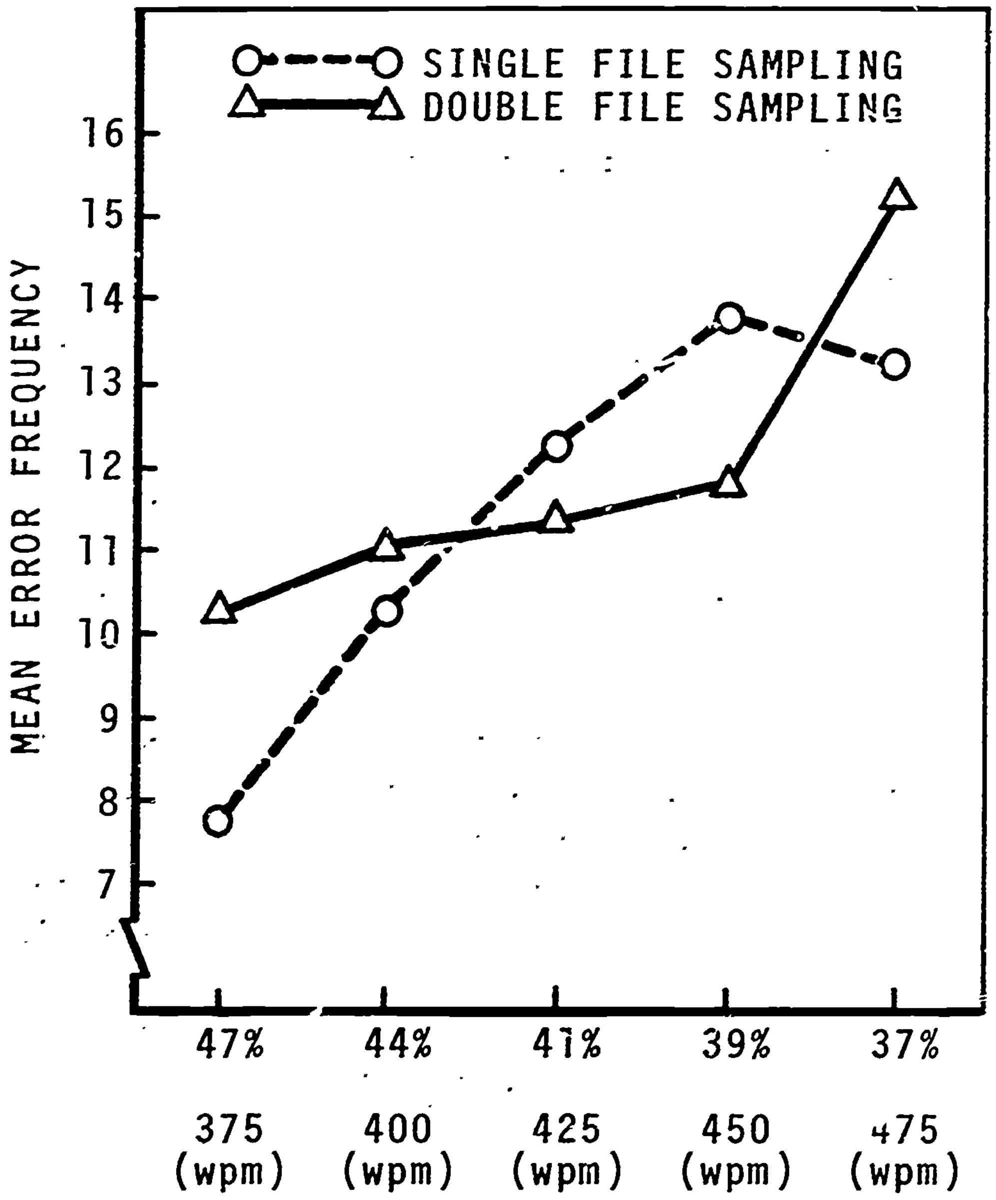

Figure 3.1 Identification Errors as a Function of Compression in Time With Method of Compression as the Parameter

TABLE 3.1

IDENTIFICATION ERRORS FOR WORDS COMPRESSED BY SINGLE AND DOUBLE FILE SAMPLING

Percent of Original Production Time	Method of Compression			
	Single File Sampling		Double File Sampling	
	Mean # of Errors	SD	Mean # of Errors	SD
47%	7.92	2.80	10.25	2.81
44%	10.25	3.59	10.92	3.97
41%	12.33	2.53	11.25	3.63
39%	13.83	4.08	11.75	3.00
37%	13.25	4.17	15.17	3.34

of the five compressions represented in the experiment, is scaled on the *x*-axis. Fractions are expressed as percents. The entry recorded below each scaled value on the *x*-axis is the word rate that would result if a listening selection, read at the average oral reading rate of 175 wpm, were reproduced in the fraction of original production time indicated by that value. The *y*-axis is scaled in terms of error scores. This figure indicates an orderly growth in error scores as the fraction of original production time required for compressed reproduction is reduced. On the other hand, the differences associated with the methods of compression appear to be small and unsystematic.

The apparent outcome of the experiment was checked by an analysis of variance of error scores, with scores classified according to amount of compression and method of compression, and with repeated measures on the methods variable. The results of this analysis are shown in Table 3.2. The growth in errors accompanying the reduction of time available for compressed reproduction was significant at the .01 level, but the variance associated with the method of compression did not reach significance at the .05 level. The interaction between these variables was significant at the .05 level.

A test of simple main effects was made in order to examine the influence of method more closely. The results of this analysis are shown in Table 3.3. The significant fact recorded in this table is

TABLE 3.2

ANALYSIS OF VARIANCE OF

IDENTIFICATION ERRORS

Source of Variation	df	MS	F
Level of Compression	4	93.[illegible]	5.26**
Error (between)	55	17.76	
Method of Compression	1	3.68	0.46
Level X Method of Compression	4	21.70	2.71*
Error (within)	55	8.00	

*p<.05
**p<.01

TABLE 3.3

ANALYSIS OF VARIANCE OF SIMPLE

MAIN EFFECTS

Source of Variation	df	MS	F
Method of Compression for 375 wpm	1	32.67	4.08*
Method of Compression for 400 wpm	1	2.67	0.33
Method of Compression for 425 wpm	1	7.04	0.88
Method of Compression for 450 wpm	1	26.04	3.25
Method of Compression for 475 wpm	1	22.04	2.75
Error	55	8.00	

*p<.05

that differences in error scores as a consequence of the method of compression used were not significant except for those words compressed to 47% of original production time.

The Newman-Keuls Test for Ordered Pairs of Means was performed in order to determine the effect of compression more precisely. Since differences due to method were, with one exception, not significant, the error scores obtained at each fraction of original production time were pooled. The results of this analysis are shown in Table 3.4.

TABLE 3.4

NEWMAN-KEULS TEST FOR ORDERED PAIRS OF MEANS

Fraction of Original Production Time	47%	44%	41%	39%	37%
47%	47%	44%	41%		
44%		44%	41%	39%	
41%			41%	39%	37%
39%				39%	37%
37%					37%

This table is arranged in matrix form, with the fractions of original production time in which words were reproduced displayed in decreasing order along the top, and down the left hand margin of the table. Entered in each row, under the appropriate column headings, are the fractions of original production time for which error scores were not significantly different from the error score associated with the fraction of original production time, recorded in the left hand margin, which identifies that row. If the table is examined as a whole, the effect of the compression variable is depicted by the total array of entries in the table.

Discussion

A significant interaction between method and amount of compression would be an interesting finding. However, since the general effect of varying the method of compression was not statistically significant, and since the differences at the various fractions of original production time were unsystematic and insignificant with one exception, the interaction that was found in the present experiment is probably without experimental significance. Where it was observed, the difference in favor of dichotic speech was probably the accidental result of uncontrolled factors in the experiment, such as differences in the recording quality of the tape bearing the words used in this comparison, or a higher frequency of sampling accidents in the 50 words processed by the electromechanical compressor.

The intelligibility of words compressed by double file sampling has been compared with the intelligibility of words compressed by single file sampling in an experiment reported by Gerber (1968). His results cannot be directly compared with the results of the present experiment, since the words he used for testing were reproduced in 50% of original production time or more, while the words used in the present experiment were reproduced in less than 50% of original production time. In Gerber's experiment, words were compressed to 75%, 67%, and 50% of original production time and, at each compression, samples with durations of 30, 40, and 50 msec. were discarded. In all of the nine comparisons provided by his experiment, he found a difference in favor of dichotic presentation. When the discarded samples were 50 msec. in duration, this difference was significant at all three compressions. However, in the six comparisons in which the discarded samples were 30 and 40 msec. in duration, three of the differences were statistically insignificant, and the remaining three, though significant, were relatively small.

The fact that Gerber found a consistent difference in favor of dichotic presentation, when the discarded samples were 30 and 40 msec. in duration, while the present experiment revealed no consistent advantage for dichotic presentation, may be, in part, a consequence of differences in the range of the compression variable explored by the two experiments. Since, in Gerber's experiment, none of the words were reproduced in less than 50% of original production time, dichotic presentation preserved all of the original speech signal. Since, in the present experiment, all the words were reproduced in less than 50% of original production time, dichotic presentation did not completely eliminate the necessity of discarding some of the speech signal. Even though discarded samples are quite small when double file sampling and dichotic presentation are used to reproduce words in less than 50% of original production time, sampling accidents are still possible, and may have injured the intelligibility of some of the words that were presented dichotically in the present experiment.

Though Gerber feels that his experiment has demonstrated the superiority of dichotic presentation, it seems to this writer that the differences he found, even when statistically significant, were too small to be of practical significance, except when the discarded samples were 50 msec. in duration. Of course, when speech is compressed by single file sampling, and when discarded samples are 50 msec. in duration, it is probable that some of the critical features of speech signals will fall entirely within discarded samples. If single file sampling is to be successful, the discarded samples must be kept short enough so that every critical feature of a speech signal has the opportunity to be sampled. As Garvey has shown (1953b), this condition is met fairly well when the discarded samples are no longer than 40 msec. in duration. In general, it can be said that the intelligibility of words is preserved better by double file sampling than by single file sampling when the discarded samples are long enough so that some of the critical features of speech signals can fall entirely within discarded samples, but that as the duration of discarded samples is shortened, the superiority of double file sampling is diminished. The results of both Gerber's experiment and the present experiment suggest that at 40 msec., this superiority has nearly vanished. Though the experience of listeners, and the examination of spectrographic records (see pg. 132, ln. 36), suggests that critical features of the speech signal may occasionally be insufficiently sampled when the discarded samples are 40 msec. in duration, the effects of such sampling accidents are counteracted by other factors, such as the listener's knowledge of the sequential dependencies inherent in sequences of phonemes and syllables.

ERIC

CHAPTER IV

REACTION TIME AS AN INDEX OF THE INTELLIGIBILITY OF TIME COMPRESSED WORDS

by

Emerson Foulke

Abstract

An experiment was performed in which three common, monosyllabic, rhyming words, compressed in time to various fractions of their original production time by the sampling method, were presented to listeners, and RT, or the time required for their identification, was determined. Reaction time decreased as word duration was decreased until a compression of 64% of original production time was reached, but was unaffected by further decreases in word duration. An effort was made to relate these results to the results typically observed in studies of listening comprehension as a function of word rate.

In evaluating the ability of listeners to process time compressed speech, two general approaches have been taken. In one approach, an effort is made to determine the intelligibility of single time compressed words or short sequences of time compressed words. In the other approach, an effort is made to determine the listener's ability to comprehend time compressed connected discourse. Word intelligibility must be one of the factors influencing listening comprehension. However, the apparent finding that word intelligibility is degraded much less by compression in time than listening comprehension (see pg. 49, ln. 6), and the finding that word intelligibility can be substantially degraded without affecting listening comprehension (Foulke, see pg. 49; Sticht, 1969; French & Steinberg, 1947), suggests that other factors must also influence listening comprehension. However, there are problems associated with the measurement of both word intelligibility and listening comprehension, and before examining this question further, it may be necessary to inquire more carefully into the operations that define both measures.

ERIC

This report is concerned with the measurement of word intelligibility. In the typical approach toward the assessment of word intelligibility, the behavior of a listener, who is instructed to reproduce a heard word, provides the evidence for intelligibility. If the listener's reproduction is accurate, it is concluded that the word was intelligible to him. If he reports on a series of such words, either the fraction he reproduces accurately, or the fraction he misses, can be taken as an index of intelligibility. In a typical experiment involving this method of measurement, an intelligibility score is obtained for words compressed by various amounts (Garvey, 1953b; Kurtzrock, 1957; Fairbanks & Kodman, 1957).

One is probably also measuring intelligibility when the ability of a listener to reproduce groups of words, such as phrases of sentences, is assessed. However, whereas the intelligibility of a single word is primarily a function of the characteristics of the speech signal, the cues that are available to a listener who knows about the sequential dependencies inherent in his language and something about the semantic import of what he is hearing, play a large part in determining the intelligibility of phrases and sentences. As the length of a sentence is increased, a point is reached at which the listener can no longer hold in storage the words, in proper sequence, he has heard. At this point, if he is to report on what he has heard, he must construct a gist recall that preserves the meaning, but not the exact form of the stimulus material. This process is much more complex than the process underlying the behavior that constitutes the evidence for word intelligibility, and it is the process upon which listening comprehension depends.

When word intelligibility is measured in the manner so far described, the listener is usually given ample time in which to reproduce each of the words he hears. However, the intelligibility that counts, if one is interested in the relationship between word intelligibility and listening comprehension, is the intelligibility of a word that occurs as a part of a continuously accumulating input that must be continuously processed by the listener. As he listens to connected discourse, he does not have the time for a leisurely and deliberate consideration of his uncertainty regarding a particular word. He must deal with incoming words quickly, and perform the selection, simplification, reorganization, or whatever encoding processes are required to transduce the information contained in the incoming speech to a form suitable for the long term storage upon which the behavior that constitutes the evidence for listening comprehension depends. Therefore, in assessing word intelligibility, it may be necessary to know not only what word the listener reproduces upon hearing a word, but also the time he requires in order to achieve that reproduction. For instance, suppose that a listener correctly identified two heard words, and that he required one-half second for the identification of one word, and five seconds for the identification of the other word. If accuracy of identification were the only evidence considered, it would be concluded that the two words were equally intelligible. And yet, if the word requiring five seconds for identification had occurred in a context of connected discourse, either it would have been unintelligible, or else the listener would have had to ignore subsequent words while attending to its identification. In terms of this analysis, it follows that the consideration of the time required for the identification of a word, in addition to the accuracy of its identification, should permit a more sensitive assessment of word intelligibility. Accordingly, an experiment was performed in which RT, the time required for the identification of a heard word, was measured as a function of the amount of compression in time.

Method

Subjects

Thirty-six students, enrolled in an introductory psychology class at the University of Louisville, served as Ss. There were 21 males and 15 females, all of whom were free from obvious hearing defects. All Ss were unfamiliar with the procedure followed in RT experiments, without experience in listening to compressed speech, and unaware of the purpose of this experiment.

Apparatus and Materials

Since the purpose of the experiment was to detect differences in reaction time as a function of the amount of compression in time, an effort was made to eliminate other sources of difference, such as variations in a S's uncertainty about the words he hears, and differences in the difficulty of word pronunciation. Therefore, three familiar, monosyllabic words were chosen, and each S was acquainted with them in advance of the experiment. It was felt that the words used should be discriminable when reproduced without compression, but not so easily discriminated that their identification would present no challenge to a listener, even when compressed. Accordingly, words were chosen that rhymed, and that were different only with respect to their initial consonants. The three words were "pie", "tie", and "lie".

These words were pronounced by a professional male announcer, whose speech was recorded on tape. This "master tape" was reproduced six times on a speech compressor of the Fairbanks type (see pg. 25, ln. 4), built at the University of Louisville -- in 100%, 78%, 64%, 54%, 47%, and 41% of original production time. If fluent speech, produced at the average oral reading rate of 175 wpm (see pg. 106, ln. 2), were reproduced in these fractions of the original production time, the resulting word rates would be 175, 225, 275, 325, 375, and 425 wpm. The output of the speech compressor was recorded on tape, and this tape was cut into segments, with each segment containing one of the words reproduced on the speech compressor. These segments were then reproduced, one at a time, in five different random orders, and the output of the tape reproducer employed for this purpose was recorded on the tape used in the experiment. The time elapsing between consecutive words recorded on the tape was approximately seven seconds, but this time was varied randomly by small amounts, from word to word, in order to suppress temporal response sets.

The experimental tape was reproduced on a Viking tape recorder, model RP 61. The signal from the tape recorder was amplified by an Eico amplifier, model HF 32, and distributed to E's monitor speaker and to S's earphones. Subject sat in an IAC audiometric testing booth, model 400. The earphones, Western Electric type ANB-H-1, were fitted with circumaural ear cushions. They were obtained from military surplus.

An auxiliary output on the Eico amplifier, intended for use as a tape recorder feed, was connected to another amplifier. The output of this amplifier was rectified and applied to the coil of a relay. With this arrangement, when a word recorded on the experimental tape was reproduced, the resulting signal closed the relay, which latched, and started a Hunter Klockounter. Subject operated a keyboard with three response keys labeled "pie", "lie", and "tie". When the key corresponding to the word heard on the earphones was pressed, the latch on the relay was broken, allowing it to return to its resting state, and the Hunter Klockounter was stopped. The remaining two keys were inactive. The active key was selected by a rotary switch, operated by E. Experimenter, seated outside the testing booth, communicated with S by means of an intercommunication system.

Procedure

Subject was acquainted with the operation of the keyboard and was told that, upon hearing a word in the earphones, he was to press the corresponding key. He was requested to strive for both speed and accuracy in selecting his response. His response, and the time required for its production, were recorded by E.

Results

Since 15 reactions to words reproduced in a given fraction of original production time were obtained from each of the 36 Ss, there were 540 observations of RT at each of the six compressions represented in the experiment. The means and standard deviations of these RTs are shown in Table 4. 1.

TABLE 4. 1

MEANS AND STANDARD DEVIATIONS OF REACTION TIME FOR TIME COMPRESSED WORDS

Fraction of Original Production Time	M	SD
100%	418 msec.	169 msec.
78%	409 msec.	160 msec.
64%	398 msec.	163 msec.
54%	403 msec.	156 msec.
47%	403 msec.	155 msec.
41%	403 msec.	151 msec.

The standard deviations recorded in column 3 suggest considerable variability of RT.

The means recorded in column 2 of Table 4. 1 were used in plotting the curve in Figure 4. 1. The scale values on the x-axis of this figure

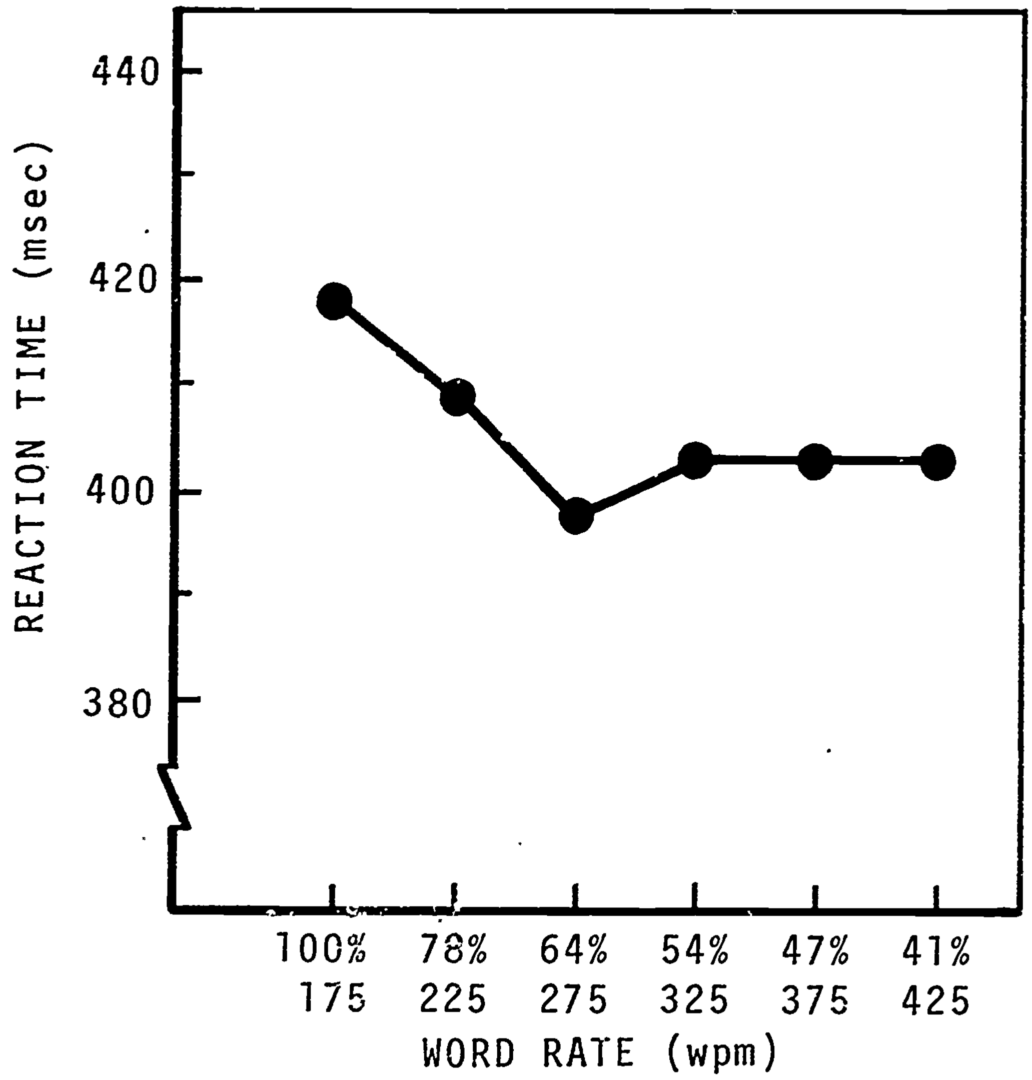

Figure 4.1 The Mean RTs for 36 Ss to Verbal Stimuli Presented at Six Levels of Acceleration (words per minute - wpm)

are percentages that indicate the fractions of original production time at which words were reproduced. The number recorded beneath each percentage indicates the word rate that would result if fluent speech, produced at the average oral reading rate of 175 wpm, were reproduced in th.t f:action of the original production time. The y-axis is scaled in msec. This curve indicates that as the time allowed for the compressed reproduction of words is decreased, after an initial decrease in the RT associated with their identification, there is no further change.

The data were examined by a Friedman two-way analysis of variance of RT (Siegel, 1956, pp. 156-172). This analysis indicated that differences associated with the changes in the time allowed for the compressed reproduction of words were not significant at the .05 level.

Discussion

The outcome of this experiment was, of course, contrary to expectations. Up to a point, reproducing words in less than the original production time seemed to have the effect of increasing, rather than decreasing their discriminability. Though further reductions in reproduction time did not result in further improvements in discriminability, neither did they result in decreased discriminability.

This experiment was preliminary in character, and was intended to probe a new avenue of research. Its outcome was too tentative to support definite conclusions. In subsequent research, experiments must be performed in which the number, structure, and familiarity of words involved in the choice is varied. The use of practiced Ss might further reduce intrasubject variability. However, in spite of the limitations of this experiment, it did hint at a relationship between the amount by which words are compressed, and the time required for their identification.

Furthermore, such a relationship, if it can be confirmed, is reasonable in view of the results that are usually obtained when listening comprehension is measured as a function of compression in time. These studies (Fairbanks, et al., 1957a; Foulke, et al., 1962; Foulke, 1968; Reid, 1968) are in general agreement regarding the finding that increasing the word rate has little effect on listening comprehension until a word rate in the neighborhood of 275 or 300 wpm is reached, but a marked effect thereafter. If the rate at which words occur in fluent speech is increased, less time will be available for the identification of words. If the listener's speech processing rate is to keep pace with an increased input rate, he must identify words more rapidly than he does at a normal rate. If, as word rate is increased and word duration is shortened, there is a point beyond which the time required by the listener to identify words is not further reduced, the result will be an insufficiency of time in which to identify words. There will be an accumulation of unprocessed input and, when the capacity for storing unprocessed input has been exceeded, listening comprehension must decline. In the present study, there was a suggestion that the time required for the identification of words decreased as their durations were decreased, until they were compressed to 64% of original production time, but not thereafter. Two hundred seventy-five wpm is the approximate word rate beyond which listening comprehension begins to decline rapidly, and 275 wpm is the word rate that results when fluent speech, recorded at the average oral reading rate of 175 wpm, is compressed to 64% of original production time.

CHAPTER V

THE INTELLIGIBILITY AND COMPREHENSION OF TIME COMPRESSED SPEECH*

by

Emerson Foulke and

Thomas G. Sticht

Abstract

A listening passage and a list of phonetically balanced (PB) words were presented at five compressions in time: 22%, 36%, 46%, 53%, and 59%. Compression was accomplished by a method which avoids distortions in vocal pitch and quality. Listening comprehension and word intelligibility were measured at each of the five time compressions. The results showed that, although both intelligibility and comprehension decreased as the percent of compression was increased, comprehension declined much more rapidly than intelligibility. An interpretation of the results is given in terms of the differential perceptual and cognitive tasks confronting the listener in the comprehension and intelligibility procedures.

Time compressed speech is speech that is reproduced in less time than the time required for the original recording. A familiar method for accomplishing this is the reproduction of a record or tape at a faster speed than the one used during recording. However, this method produces distortion in vocal pitch and quality that interfere seriously with its intelligibility.

---

*An account of the research reported in this chapter can also be found in the Proceedings of the Louisville Conference on Time Compressed Speech, Louisville: University of Louisville, 1967, 21-28. The authors wish to express appreciation for the helpful comments of Dr. Doris Aaronson, Center for Cognitive Studies, Harvard University, who read the manuscript.

ERIC

Speech may also be compressed in time, and without distortion in vocal pitch, by a sampling method in which brief segments of recorded speech are periodically discarded and the resulting gaps are closed. The success of the sampling method depends upon the fact that samples can be discarded which are so small that the human ear cannot detect their absence.

Compression of this sort may be accomplished manually by removing short segments of a recorded tape and splicing the free ends together again (Garvey, 1953b). If, for instance, every third centimeter of a recorded tape were removed in this manner, the resulting tape would be two-thirds the length of the original tape, and only two-thirds as much time would be required for its reproduction.

The manual sampling method is, of course, too cumbersome for most purposes. Equipment utilizing a method introduced by Fairbanks, et al., (1954) accomplishes a similar kind of compression by electro-mechanical means.

The superiority of the sampling method with respect to the intelligibility of single words has been demonstrated by Garvey. He compared the intelligibility of words compressed in time both by the sampling method and by increasing the playback speed of recorded tape, and found that listeners could identify a significantly higher percentage of words compressed in time by the sampling method.

The superiority of the sampling method cannot be demonstrated so easily when the listener's task is changed from mere identification of words, as in the intelligibility testing procedure, to the comprehension of connected speech. Foulke, et al., (1962), found substantial losses in the comprehension of listening selections, as indicated by performance on multiple-choice tests, when the selections were compressed enough to produce word rates in excess of 275 wpm. Thus, it appears that compressions that interfere very little with intelligibility, interfere substantially with comprehension.

In a direct comparison of a listening selection compressed both by the sampling method and by increasing the playback speed of tape, McLain (1962) found a slight but statistically significant difference in favor of the sampling method for a selection reproduced at 325 wpm. Foulke (1966a), in an experiment that presented a listening selection compressed by both methods, and at several accelerated word rates, found no differences in comprehension that could be attributed to the methods of compression.

The foregoing evidence, though scattered, suggests that connected discourse which has been compressed in time may not be comprehensible, even though the individual words in such discourse remain intelligible when presented at the same compression. However, there has been no single experiment in which intelligibility and comprehension have been examined over a wide range of compressions in time. The issue at stake here is an important one since a definitive answer to the question has important implications for future research. To the extent that the problem is one of loss of intelligibility of single words, attention will be directed toward the improvement of the equipment used for time compression. To the extent that the problem is the increased rate at which information is fed to the central nervous system when speech is compressed in time, attention will be directed to the analysis of the demands placed upon the perceptual and cognitive processing functions of the listener by time compressed speech. Because of these considerations, an experiment was performed in which the intelligibility of single words and the comprehension of connected speech were measured at several compressions in time.

Method

Subjects

One hundred University of Louisville students, of both sexes, served as Ss in the experiment. All were free from any obvious hearing defects and none of them had prior experience with time compressed speech.

Apparatus and Materials

Listening comprehension was measured with the listening subtest of the Sequential Test of Educational Progress, Form 1A, Part 1. Form 1A consists of brief listening selections of scientific and literary content that are appropriate with respect to interest and difficulty for a college freshman population. For each selection, there are a few multiple-choice questions covering facts and implications of the selection. Part 1 contains five such selections and a total of 36 questions. Due to an inadvertance, question 17 was omitted, so that the highest possible test score in the present study was 35.

The five listening selections were read in a recording studio at the American Printing House for the Blind by a professional reader employed in the Talking Book program, and were recorded on magnetic tape by an Ampex tape recorder, model 300. This tape was then compressed in time by means of the Tempo Regulator, a device that accomplishes compression by Fairbanks' sampling method discussed earlier*.

The master tape, recorded at a word rate of 175 wpm, was reproduced on the Tempo Regulator at those compressions required to produce word rates of 225, 275, 325, 375, and 425 wpm. The output of the Tempo Regulator was recorded on magnetic tape and this tape was reproduced, during the experiment, on a Wollensak tape recorder, model T-1500. The output of the tape recorder was distributed to the Ss through headsets fitted with ear cushions, and the signal level at each headset could be adjusted by the S for comfortable listening.

The 100 words comprising a phonetically balanced word list were read by the same reader, prepared in the same manner, and compressed on the Tempo Regulator by the same percentages as the listening selections (Egan, 1948). As before, the output of the Tempo Regulator was recorded on tape and this tape was used in the experiment.

Finally, a brief "warm up" listening selection was prepared at each of the compressions represented in the experiment. This selection was used to promote a common listening set by providing Ss with brief experience in listening to time compressed speech before participating in the experiment.

Procedure

The 100 Ss were distributed among 5, 20 member groups. Each group heard material reproduced at one of the compressions used in the experiment. All of the members in each group listened to the "warm up" passage first. Then, each group was further divided into two sub-groups. The members of one sub-group heard and were tested on the listening selections first and then identified, in writing, the phonetically balanced words, which were presented one at a time with a five second interval between words. This order was reversed for the

*For further information about speech compression equipment, consult Infotronic Systems, Inc., 2 West 46th Street, New York, New York 10063. Readers interested in obtaining time compressed tapes for research or demonstration may write to Dr. Emerson Foulke, Director, Center for Rate Controlled Recordings, University of Louisville, Louisville, Kentucky 40208.

other sub-group, to control for the possibility of an effect due to order. The same Ss were used for the measurement of intelligibility and of comprehension in order to suppress effects due to individual differences.

Subjects were tested as they became available. Therefore, although several Ss were usually tested at a time, occasionally only one S was present at a testing session. Tests were conducted at a given compression until the 20 Ss required for an experimental group had been tested. This procedure was followed for the five experimental groups.

Results

An intelligibility score, the percent of correctly identified PB words, and a comprehension score, the percent of correctly answered standard deviations of these scores at each of the five time compressions represented in the experiment are shown in Table 5.1. The effect of

TABLE 5.1

CHANGES IN INTELLIGIBILITY AND COMPREHENSION AS A FUNCTION OF PERCENT OF COMPRESSION IN TIME

Percent of Compression	Intelligibility		Comprehension	
	Mean	SD	Mean	SD
22%	93%	2.2	73%	12.4
36%	91%	3.0	66%	14.7
46%	89%	3.2	67%	13.0
53%	85%	5.0	56%	12.0
59%	84%	3.7	53%	14.0

time compression on intelligibility and comprehension is also shown in Figure 5.1. In this figure, the five time compressions employed in the experiment are displayed along the x-axis. The entry below

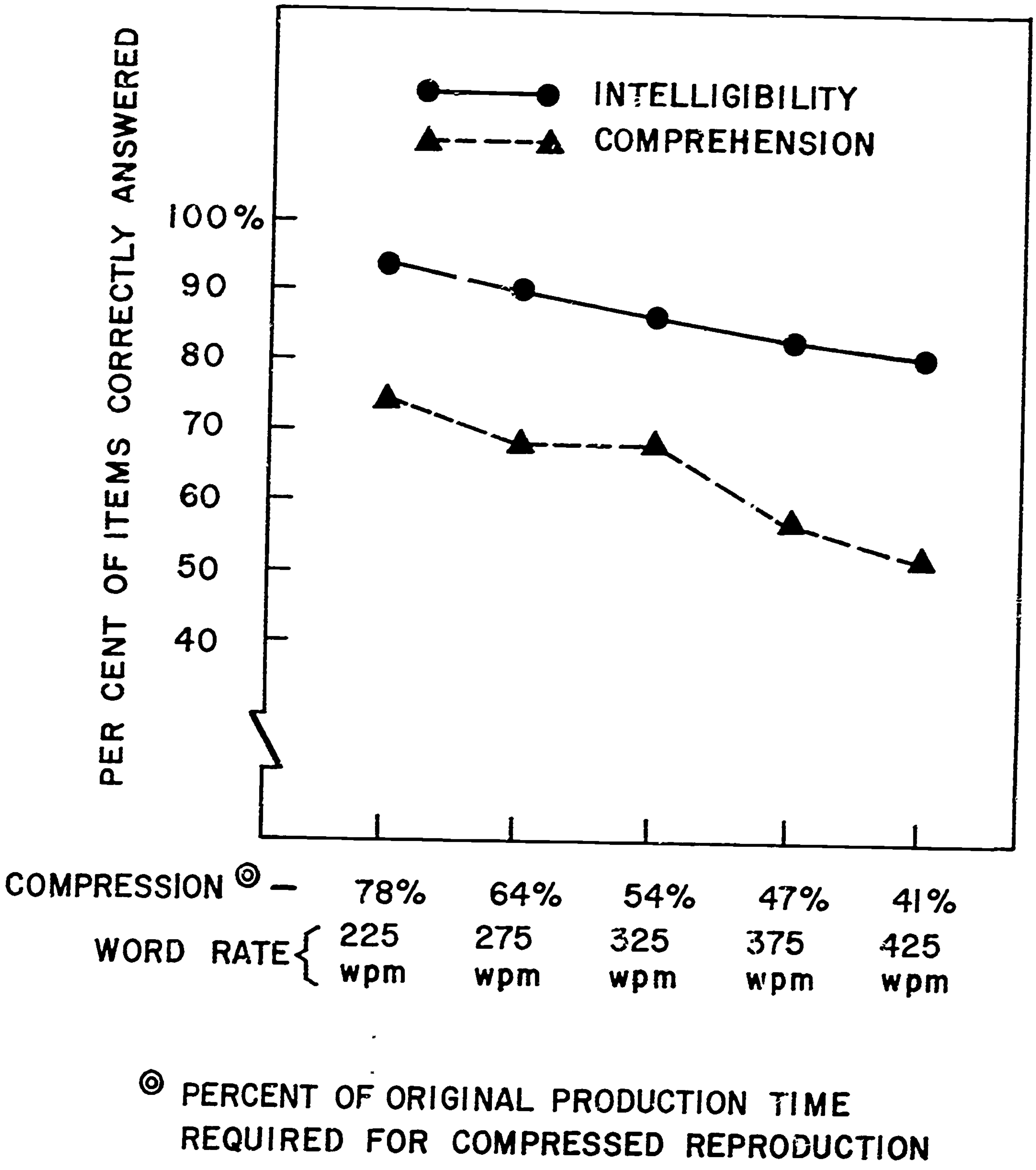

Figure 5.1 Word Intelligibility and Listening Comprehension as a Function of Percent of Compression

each compression value refers to the word rate that would result if connected discourse at a normal word rate of 175 wpm were compressed by that amount (Johnson, et al., 1963). Percent correct for the two dependent variables is scaled on the y-axis. As the amount of compression was increased, both intelligibility and comprehension decreased. However, comparison of the two curves indicates that intelligibility was always superior to comprehension and that intelligibility was affected much less than comprehension by increasing the amount of compression*.

The data upon which Figure 5.1 is based were examined by an analysis of variance. The results of this analysis, presented in Table 5.2, confirm the impressions conveyed by Figure 5.1. Changes

TABLE 5.2

THE ANALYSIS OF VARIANCE OF INTELLIGIBILITY

SCORES AND COMPREHENSION SCORES

Source	df	M	F
Between Ss	99		
Percent of Compression	4	1,449	15*
Error (b)	95	99	
Within Ss	100		
Intelligibility vs. Comprehension	1	32,462	877*
Interaction	4	891	6*
Error (w)	95	37	

*p < .001

*A graph, like the graph in Figure 5.1, was constructed, using intelligibility and comprehension scores that had been corrected for guessing by the same formula. The difference between the relationships depicted between the two curves in this graph were more in intelligibility and in comprehension, as well as the interaction of these variables, were significant ($p < .001$ in all cases).

Discussion

With respect to intelligibility, the results of the present study are in good agreement with those of Garvey. There was only a 9% loss in the intelligibility of PB words compressed by an amount sufficient to produce a word rate of 425 wpm with connected speech, assuming an original or uncompressed word rate of 175 wpm. At the compression that would be required to accelerate speech to approximately twice the normal word rate, there was only a 6% loss in the intelligibility of PB words. At a similar compression accomplished by the alternative method of reproducing a tape at a faster speed than the one used during recording, Klumpp and Webster reported a 60% loss in intelligibility (Klumpp & Webster, 1961). Garvey also found intelligibility losses of this magnitude when compression was accomplished by increasing the playback speed of tape. Thus, we conclude with Garvey that the intelligibility of single words is affected much less by the sampling method than by the speeded playback of a tape or record. The superiority of the sampling method in this respect is probably explained adequately by its freedom from distortion in vocal pitch and quality.

It was, of course, expected that comprehension scores would be lower than intelligibility scores. The demonstration of comprehension imposes a much more complex task on the listener than does the demonstration of intelligibility. The behavior upon which the measurement of intelligibility depends, implies registration of the stimulus word, some kind of short term memory storage, and the transduction of the stored item to an overt response. On the other hand, the behavior on which the measurement of comprehension is based, implies continuous registration and short term memory storage

---

pronounced than the difference suggested in Figure 5.1. If the formula used to correct intelligibility scores for guessing had reflected the very small probability of choosing the correct answer by chance, the difference between the two curves would have been even greater. For these reasons, uncorrected scores were used in Figure 5.1 and the analysis reported in Table 5.2, because this seemed to be a more conservative course.

of stimulus material, the continuous encoding, or simplification by reorganization and selective discarding of stimulus information so that it can be transferred to long term memory storage, and a final decoding step required for the transduction of material in long term storage to overt behavior.

It is the finding that the difference between intelligibility and comprehension scores increases as the amount of compression is increased that requires additional explanation. One possibility is that the progressively larger loss in comprehension is a consequence of the cumulative effects of the relatively smaller losses in intelligibility. The data of the experiment were examined for this possibility in the following manner. All of the Ss tested at a given compression were separated into a high and a low scoring group, on the basis of their comprehension tests scores. The difference between the means of the intelligibility scores of the two groups formed in this manner, was tested for significance. In all but one case, (the 59% compression group) the difference between means did not reach significance at the 5% level. This finding suggests that, with respect to the results of the present experiment, poor comprehension cannot be satisfactorily explained by low intelligibility for individual words. In any case, it is well known that it is not necessary for all of the units of a message to be intelligible in order for the message to be received accurately (Miller & Selfridge, 1950; Attneave, 1954). Because of prior learning, the listener is able to reconstruct a sent message on the basis of reduced cues. He makes use of sequential probabilities in grammatical speech and the meaningfulness of the heard message in supplying missed words.

A more convincing explanation may be that when continuous speech is compressed, the number of words per unit time is increased, and the intervals between words are decreased. It has been shown repeatedly in studies of verbal learning that the difficulty of a learning task is increased by increasing the number of items in the list to be learned and by decreasing the interstimulus interval (Miller, 1951; Osgood, 1953; Aaronson, 1968). To the extent that these two situations are similar, an increase in time compression may mean an increased contribution of factors related to task difficulty. Such factors would not apply to the measurement of intelligibility, as defined in this study, since its measurement required the presentation of single words in isolation, rather than connected sequences of words.

The results of the present study suggest the relevance of a concept such as channel capacity (Miller, 1953, 1956). According to this concept, a communication channel, in this case the listener, has a finite capacity for handling information. As the amount of information applied to the input of the channel is increased, there is a corresponding increase in the amount of information transmitted by the channel, until channel capacity is reached. Further increases in the amount of input information cannot be handled by the channel, with the result that some information is lost. Assuming normal speech to occur at a rate that is well below channel capacity, increasing word rate should have little effect upon comprehension initially. However, as the word rate reaches channel capacity comprehension should begin to decline, and, when channel capacity has been exceeded, comprehension should fall off very rapidly. The comprehension curve in Figure 5.1 resembles a positively accelerated decreasing function, although not enough values for the word rate variable were determined to test this suggestion. However, the results of other studies have also suggested that comprehension is a positively accelerated decreasing function of word rate (Foulke, 1964a).

Silent visual reading rates considerably in excess of 275 wpm, the word rate at which listening comprehension generally begins to decline rapidly, are commonplace. However, because of the spatial display of information on the printed page, the reader is able to perform the perceptual operation referred to by Miller as "chunking". In order to keep the rate of information input below his channel capacity, the fast visual reader reduces the number of elements with which he must contend by combining the elements given by the structure of language into larger elements. He begins to perceive not just single words, but entire phrases or sentences. Because of the temporal display of information presented aurally, the listener cannot perform this operation.

The data required to test the explanation offered here are not yet available. One clear task for future research is a more careful determination of the relationship between word rate and comprehension. If, after further investigation, the attempt to determine the differential effect of increasing word rate on intelligibility and comprehension of compressed speech is convincing, it will have important practical implications. If the inability to show good comprehension of very rapid speech is found to be a consequence of a verbal input that has been rendered incompatible with the human perceptual mechanism because channel capacity has been exceeded, current efforts to train for comprehension of very rapid speech cannot be expected to have much effect. This conclusion is not contradicted by past efforts at training. Such efforts have not, in the main, been successful (Voor & Miller, 1965). However, the task of defining an adequate training experience has only begun, and further efforts along this line are now in progress (Orr, et al., 1965).

If, on the other hand, loss in comprehension turns out to be primarily a consequence of words that are less intelligible because of the degradation of signal quality that is inherent in the time compression of speech by the sampling method, other directions for research are indicated. For instance, one might consider further engineering refinements of the equipment used for the time compression of speech, with a view to improving signal quality. One might also consider a training program designed to promote the comprehension of highly compressed continuous speech by teaching listeners to discriminate and identify words and phrases that are rendered unfamiliar by virtue of having been greatly compressed in time.

CHAPTER VI

LISTENING COMPREHENSION AS A FUNCTION OF WORD INTELLIGIBILITY

by

Emerson Foulke

Abstract

An experiment was performed in which five versions of a recorded listening selection, differing systematically with respect to vocal pitch, were compressed to 54% of the original production time. The reader's normal vocal pitch was the lowest of five pitches used. Pitch was increased, from version to version, in equal steps, through a range of approximately one octave. Research has shown that the intelligibility of words compressed in time by a sampling method that preserves vocal pitch is not seriously affected until an extreme compression is reached, but that when words are reproduced by a method which produces pitch distortion, intelligibility is seriously affected. Since the five listening selections in this experiment were different with respect to vocal pitch, there should have been differences in the intelligibility of the words with which they were composed. If listening comprehension is a function of word intelligibility, this fact should be reflected in the comprehension test scores of Ss who listened to the five versions of the selection. Each of the five versions was presented to a different one of five comparable groups of Ss, who were then

ERIC

tested for listening comprehension. There were no significant differences in comprehension related to the pitch at which the listening selection was reproduced, suggesting that listening comprehension was not affected by the variations in word intelligibility produced by this method.

A spoken word is intelligible if, when presented in isolation, it can be reproduced accurately by a listener. Comprehension is revealed by the ability to demonstrate knowledge of the facts and implications of a listening selection. The behavior that constitutes the evidence for word intelligibility requires only the short term storage of a stimulus item necessary for immediate recall. The behavior that constitutes the evidence for comprehension requires, in addition, encoding and decoding processes, and long term storage.

There is, of course, a relationship between word intelligibility and listening comprehension. If the individual words of a listening selection were completely unintelligible, the listener could not comprehend the listening selection. However, there are reasons to believe that a point is reached beyond which further improvements in the intelligibility of the words in a listening selection will not result in further gain in listening comprehension.

Garvey (1953b) compared the intelligibility of words compressed in time by reproducing a tape at a faster speed than the one used during recording with words compressed in time by a sampling procedure in which brief segments of the recorded tape were regularly eliminated. The first method results in an elevation of vocal pitch that is proportional to the increase in playback speed of the recorded tape. The second method leaves the pitch of the speaker's voice undisturbed. When, by increasing tape playback speed, words were reproduced in 50% of the time required for original production, there was a 35% loss in intelligibility, and a 92% loss in intelligibility when they were reproduced in 40% of the original production time. On the other hand, when, by the sampling method, words were reproduced in 50% of the original production time, there was only a 5% loss in intelligibility, and a 7% loss in intelligibility when they were reproduced in 40% of the original production time. The two methods of time compression have the same effect on the rate at which speech sounds occur. However, since compression by increasing tape playback speed elevates vocal pitch while compression by periodic sampling does not, it is probably the elevation in vocal pitch that is primarily responsible for the loss in intelligibility.

However, when the two methods for the time compression of speech were compared with respect to the comprehension of connected discourse, there was little or no difference between them. McLain (1962) compressed a listening selection to 54% of its original production time by each of the two methods. The comprehension test scores of two groups of Ss who had listened to these compressed selections were compared and a statistically significant but rather small difference in favor of the sampling method was found. Foulke (1966a) performed a similar experiment in which a listening selection was compressed to 70% (250 wpm), 58.33% (300 wpm), and 50% (350 wpm), of the original production time by each of the two methods. The groups of Ss who heard the six resulting versions of the listening selection were tested for listening comprehension. There was no difference in the outcome of the experiment that could be associated with the method used for time compression. Thus, the difference in favor of the sampling method, when the comparison is made in terms of word intelligibility, largely or completely disappears when the comparison is made in terms of comprehension.

It is possible, by combining the two methods for the time compression of speech, to hold constant the rate at which speech sounds occur, while varying the amount of distortion in vocal pitch. That is, if, for each of several versions of a listening selection, the two methods for time compression are combined in different proportions to produce the same final accelerated word rates, the resulting versions of the listening selection will vary with respect to distortion in vocal pitch. Since there is a strong relationship between distortion in vocal pitch and word intelligibility, this scheme provides a method for varying word intelligibility systematically. Of course, the versions resulting from this treatment will also vary with respect to the amount of speech information that has been discarded. But, as has already been shown, the sampling method has a relatively small influence on word intelligibility.

The finding that it is not necessary for all of the words in a listening selection to be intelligible in order for that selection to be comprehensible is explained by the ability of the listener to make use of the redundancy in spoken language to recover missed words or meanings (Miller & Selfridge, 1950). Klumpp and Webster (1961), for instance, report higher identification scores for time compressed phrases than for time compressed single words. However, a more systematic exploration of the relationship between word intelligibility and the comprehensibility of connected discourse would promote a better understanding of the cognitive contribution of the listener to the task of comprehending. In the experiment, the report of which follows, the word intelligibility of a listening selection has been varied systematically by varying the distortion in vocal pitch, while holding word rate constant.

Method

Subjects

One hundred sixty-one seventh, eighth, and ninth grade pupils, of both sexes, from four residential schools for the blind, served as Ss in the experiment. Subjects were assigned to five experimental groups in such a way that the proportional representation of schools and of grades was approximately the same for all groups. The five groups contained 34, 34, 29, 32, and 32 members respectively.

Experimental Materials and Apparatus

The listening selection was a 3,350 word fictional account of a boy's encounter with a band of pirates on a desert island. It was judged to be appropriate in interest and difficulty for children in the seventh, eighth, and ninth grades (Allen, 1958). This selection was read orally by a professional reader and recorded on magnetic tape by means of an Ampex tape recorder, model 300, in the Talking Book Studios of the American Printing House for the Blind.

This "master tape" was used to prepare five versions of the listening selection, each compressed to approximately 54% of its original length. This magnitude of compression was chosen because previous research (Fairbanks, et al., 1957; Foulke, et al., 1962) has shown it to be in the middle of the range in which changes in compression are accompanied by changes in listening comprehension. If word intelligibility is a factor in listening comprehension, its systematic variation should affect the comprehension of speech compressed by this amount. Version 1 was made by reproducing the "master tape" on the Tempo Regulator at the desired amount of compression. The output of the Tempo Regulator was recorded on the tape to be used in the experiment by means of a Crown tape recorder, model 800. Thus, the compressed speech in Version 1 was accomplished entirely by the sampling method, and it was free from distortion in vocal pitch. In

ERIC

Version 2, the "master tape" was reproduced on the Tempo Regulator, adjusted for three-fourths of the desired compression in time, and its output was recorded on tape by means of the Crown tape recorder. The remaining compression was accomplished by reproducing this tape at a faster speed than the one used during recording, and this speeded reproduction was recorded on tape to be used in the experiment. In Version 3, half of the desired compression was accomplished by each method. In Version 4, one-fourth of the desired compression was accomplished by the sampling method, and the remaining three-fourths by increasing tape playback speed. In Version 5, all of the compression was accomplished by increasing tape playback speed. Although the five versions of the listening selection prepared in this manner were approximately the same with respect to word rate, (approximately 325 wpm), there was a progressive elevation in vocal pitch from Version 1 to Version 5.

The tapes used in the experiment were reproduced on a Uher tape recorder, model 4000, and its output was distributed to the Western Electric headsets, type ANB-H-1, worn by the Ss. The headsets were fitted with circumaural ear cushions, and equipped with volume controls, so that the signal level could be adjusted by each S for comfortable listening.

A 42 item, four-alternative, multiple-choice test, with a split-half reliability of .76, was prepared for the listening selection. Test questions were read orally by a skilled reader, and recorded on magnetic tape by means of a Crown tape recorder, model 800, in the compressed speech laboratory at the University of Louisville. Each item, including its four alternatives, was read twice. Special answer sheets were prepared for use by blind students. For each item, the student indicated his choice of alternatives by making a pencil mark in one of four areas, outlined by braille dots and designated by braille letters.

Procedure

All of the Ss at a particular school for the blind that qualified for membership in a particular experimental group were tested at one time. First, Ss heard the tape recorded instructions for participating in the experiment, and were given practice trials in marking their answer sheets. Then, the appropriate version of the compressed listening selection was presented. Following this, the tape recorded test questions were presented and Ss marked their answer sheets. If necessary, the tape recorder was stopped between questions, until all Ss had made their choices. However, it was usually unnecessary to stop the tape recorder. This testing arrangement avoided the problem of keeping one's place, which is a serious problem for braille readers who must alternate between a question booklet and an answer sheet. It also assured that each S attempted every item on the test.

Results

Each S's score was the number of test items correctly answered. The means and standard deviations of these scores, for the five experimental groups, are shown in Table 6.1. It is clear that the

TABLE 6.1

MEANS AND STANDARD DEVIATIONS OF

COMPREHENSION TEST SCORES

Groups	I	II	III	IV	V	
n=	34	34	29	32	32	N=161
M=	18.68	18.79	19.72	19.22	18.97	
SD=	9.04	8.09	7.68	7.11	7.66	

different experimental treatments produced very little difference in mean test scores. An analysis of variance of test scores (see Table 6.2) indicated no significant differences among test scores that could be associated with experimental treatments.

Discussion

Within the range in which word intelligibility was varied in this experiment, it exerted no influence on the comprehension of connected speech. If intelligibility had been degraded sufficiently, there doubtless would have been a loss in comprehension. Nevertheless, within broad limits, listening comprehension does not appear to depend very heavily upon the intelligibility of single words. There is apparently enough redundancy in spoken language so that many words can be transmitted imperfectly, or not at all, without interfering seriously

TABLE 6.2

ANALYSIS OF VARIANCE OF COMPREHENSION TEST SCORES

Source of Variation	df	M	F
Between Groups	4	5.32	.08*
Within Groups	156	63.43	

*Not significant at the .25 level.

with listening comprehension. As a listener acquires experience with his language -- its grammar and its conventional forms -- he acquires information about the probabilities associated with the occurrence of particular words, given the occurrence of particular preceding words. Similarly, the context of meanings aroused by a listening selection reduces the listener's uncertainty, at any given instant, regarding the words and phrases that are to follow. The listener is able to use this information concerning the probabilities associated with the occurrence of words, phrases, or sentences, to reconstruct imperfectly transmitted speech.

When the outcome of this study is considered, together with the outcome of the studies in which the dependence of listening comprehension upon word rate has been investigated, it appears that listening comprehension depends more upon word rate than upon word intelligibility. If, within broad limits, listening comprehension is not markedly influenced by word intelligibility, the decline in the comprehension of speech that has been compressed in time, cannot easily be explained by the degradation of the signal imposed by the process of compression. In any case, as has already been mentioned, words can be compressed by the sampling method to less than half their original duration without a serious loss in intelligibility. The loss in comprehension at fast word rates is due not to faulty stimulus registration, but to the presentation of words at a rate that is faster than the rate at which the listener can process them.

The method employed in this experiment provides a way of investigating the contribution of the listener, with his background of experience, to the perception of spoken language. Since word intelligibility can be systematically degraded, the listener can be forced into progressively greater reliance upon his store of information regarding word probabilities in restoring imperfectly transmitted messages.

If the listener's ability to tolerate degradation of word intelligibility is explained by the redundancy in spoken language, the effect of degrading word intelligibility should depend upon the redundancy of the language to be heard. An experiment in which comprehension is determined, as a function of word intelligibility, for messages the redundancy of which has been varied by the technique reported by Miller and Selfridge (1950), should be illuminating.

CHAPTER VII

LISTENING COMPREHENSION AS A FUNCTION OF WORD RATE*

by

Emerson Foulke

Abstract

Twelve comparable groups of Ss heard a listening selection that differed, from group to group, with respect to word rate. Word rate was varied, in increments of 25 wpm, from 125 to 400 wpm, by means of the sampling method for compressing or expanding recorded speech. After listening to the selection, Ss were tested for comprehension by a multiple-choice test. Comprehension was not seriously affected by increasing word rate from 125 to 250 wpm, but it declined rapidly thereafter. The suggested explanation of these results is that time is required for the perception of words, and that as word rate is increased beyond a certain point, the perception time available to the listener becomes inadequate, and a rapid deterioration of listening comprehension commences.

If word rate is determined for a large number of samples of the oral reading of professional readers, such as radio newscasters or those who read Talking Books, considerable variability will be observed.

---

*The material in this chapter also appears as an article in The Journal of Communication, 1968, 18, No. 3, 198-206.

This variability is the consequence of differences in the nature of the material that is read, and to differences in personal reading style. However, the mean word rate will be approximately 175 wpm (Johnson, et al., 1963; Foulke, see Chapter XI, pg. 106). Recent technological developments (Fairbanks, et al., 1954; Foulke, 1964a; Scott, 1965), have made it possible to vary word rate of recorded oral reading over a wide range, either slower or faster than normal, without distortion in vocal pitch. This capability raises the possibility of presenting speech at other rates than the one at which it happens to be produced by an oral reader. On the practical side, recorded speech at a faster than normal rate can provide a needed increase in reading speed for blind people, and other people who read by listening. Recorded speech at slower than normal rates may prove to be a useful tool in promoting certain kinds of instruction, such as the learning of a foreign language. In a more theoretical vein, the ability to vary speech rate through a wide range, suggests new avenues for investigating the cognitive processes that underlie the perception of speech.

There are several studies in which comprehension has been measured as a function of word rate; but, in each of these studies, word rate has been varied through a relatively limited range. Therefore, in order to gain an impression of the influence of this variable, it has been necessary to combine the results of several studies. Within the range extending from 126 to 172 wpm, Diehl, et al., (1959), found listening comprehension to be unaffected by changes in word rate. In the range extending from 125 to 225 wpm, Nelson (1948) and Harwood (1955) found a slight, but insignificant loss in comprehension as word rate was increased. Fairbanks, et al., (1957c), found little difference in the comprehension of listening selections presented at 141, 201, and 282 wpm. Thereafter, comprehension, as indicated by percent of test questions correctly answered, declined from 58% correct at 282 wpm to 26% at 470 wpm. Foulke, et al., (1962), using both technical and literary listening selections, found comprehension to be only slightly affected by increasing word rate up to 275 wpm. However, in the range extending from 275 to 375 wpm, they found an accelerated decrease in comprehension as word rate was increased. Foulke and Sticht (1967), using the STEP Listening Test, Form 1A, found a decrease in comprehension of 6% between 225 and 325 wpm, and a decrease of 14% between 325 and 425 wpm*. The last three studies cited are in agreement regarding the finding that there is a change in the rate at

---

*Sequential Tests of Educational Progress, Cooperative Test Division, Educational Testing Service, Princeton, New Jersey, 1957.

which comprehension declines as word rate is increased. A similar relationship has also been found in many other studies in which the determination of the influence of word rate on listening comprehension was not the primary objective (Foulke, 1966b).

The purpose of the study reported in this paper is to display the way in which listening comprehension varies as word rate is varied over a wide range. It is felt that a more certain knowledge of the relationship between these variables will be useful in making decisions about the rate at which to present recorded speech, in both practical and theoretical applications.

Method

Subjects

Three hundred sixty sighted college students of both sexes, drawn from psychology and education classes at the University of Louisville, served as Ss. In a majority of instances, their service fulfilled a course requirement. Subjects were divided into 12 experimental groups, with 30 Ss per group.

Experimental Materials and Apparatus

A 2,925 word listening selection, appropriate in interest and difficulty for a college population, was chosen for use in the experiment (Durant, 1957). A 50 item, four-alternative, multiple-choice test, with a split-half reliability of .68, was written for this selection.

The selection was read orally by a professional reader and recorded on a magnetic tape by an Ampex tape recorder, model 300, in the Talking Book Studios of the American Printing House for the Blind. This "master tape" was reproduced on a modified Tempo Regulator (Foulke, 1964a), an electromechanical device for the compression or expansion of speech (Fairbanks, et al., 1954). The Tempo Regulator was adjusted for one of the word rates to be used in the experiment, and its output was recorded on magnetic tape by a Crown tape recorder, model 800. Instructions for participating in the experiment were also recorded on this tape. Twelve tape recorded versions of the listening selection were prepared in this manner, covering the range from 125 through 400 wpm in steps of 25 wpm. The tapes used in the experiment were reproduced on a Wollensak tape recorder, model T-1500. The output of the tape recorder was distributed to the Western Electric headsets, type ANB-H-1, fitted with ear cushions, and each headset was provided with a volume control so that the signal level could be adjusted by the Ss for comfortable listening.

Procedure

It was not possible to obtain the assistance of enough Ss at any one time so that a complete experimental group could be tested at one sitting. Therefore, Ss were tested in groups that ranged from 10 to 20 in number, and tests were conducted at a given word rate until the 30 Ss required for that condition of the experiment had been tested. Starting with the slowest word rate used in the experiment, the listening selection was presented to succeeding experimental groups in ascending order of word rate.

The experiment was conducted in a large university classroom, with the poor acoustical properties typical of such rooms. However, since all Ss listened by means of headsets fitted with the kind of circumaural ear cushions that completely surrounds and encloses the external ear, the listening environment was felt to be satisfactory and similar for all Ss.

First, test booklets and answer sheets were distributed. Next, Ss heard the recorded instructions for participating in the experiment. Then, the listening selection was presented. Upon its conclusion, Ss proceeded immediately to the test of listening comprehension, and upon its completion, each S turned in his test materials and quietly left the room. Each experimental session was concluded within the 50 minute class period.

Results

A corrected test score was determined for each S by applying to his raw score the formula $CS = R - [W \div (n-1)]$ when CS = corrected score, R = right answers, W = wrong answers, and n = the number of alternatives in the test item (Cronbach, 1960, p. 50). A correction for guessing was applied to raw test scores because it was felt that the assumptions underlying a correction of this sort are reasonably met when experimental group means are to be compared. The means and standard deviations of corrected test scores, for each of the 12 experimental groups, are shown in Table 7.1.

TABLE 7.1

MEANS AND STANDARD DEVIATIONS OF COMPREHENSION

TEST SCORES AS A FUNCTION OF WORD RATE

WPM	M	SD
125	44.33	12.86
150	48.71	12.97
175	44.79	14.73
200	42.39	12.79
225	47.28	15.97
250	45.05	15.52
275	37.96	14.17
300	39.11	12.74
325	30.58	17.90
350	29.87	16.18
375	23.73	14.16
400	20.27	11.20

The relationship between word rate and mean test score, expressed as a percent of the maximum possible score, is displayed graphically in Figure 7.1. Word rate is scaled on the abscissa, and test score, in percentage units, on the ordinate. Though the curve in Figure 7.1 is somewhat irregular, the relationship suggested by it is one in which comprehension is relatively unaffected by changes in word rate in the range bounded by 125 and 250 wpm. Beyond this range, however, compression declines rapidly as word rate is increased.

The test scores used in plotting Figure 7.1 were examined by an analysis of variance, the results of which are shown in Table 7.2. The variance in test scores associated with changes in word rate is significant beyond the .01 level as shown in row 1 of this table.

TABLE 7.2

ANALYSIS OF VARIANCE OF COMPREHENSION

TEST SCORES

Source of Variance	df	MS	F
Between	11	2,956.74	14.79*
Linear	(1)	.92	
Within	348	199.88	

*p <.01

The significance of the difference between ordered pairs of individual means was examined by means of the Newman-Keuls Test for Ordered Pairs of Means (Winer, 1962, p. 80). The results of this analysis are shown in Table 7.3. This table is cast in matrix form, with the word

TABLE 7.3

NEWMAN-KEULS ANALYSIS OF THE SIGNIFICANCE

OF DIFFERENCES AMONG GROUP MEANS

WPM	125	150	175	200	225	250	275	300	325	350	375	400
125	125	150	175	200	225	250		300				
150	125	150	175	200	225	250	275	300				
175	125	150	175	200	225	250	275	300				
200	125	150	175	200	225	250	275	300				
225	125	150	175	200	225	250	275	300				
250	125	150	175	200	225	250	275	300				
275		150	175	200	225	250	275	300	325	350		
300	125	150	175	200	225	250	275	300	325	350		
325							275	300	325	350		
350							275	300	325	350	375	
375									325	350	375	400
400											375	400

rates at which tests were conducted arranged down the left hand margin and across the top of the matrix in order of increasing magnitude. Entered in each row, under the appropriate column headings, are the word rates for which comprehension scores were not significantly different from the comprehension score associated with the word rate in the left hand margin that identifies the row. The results presented in Table 7.3 are in general agreement with the impression conveyed by Figure 7.1. The pattern formed by the entries in this table also depict the nature of the relation between word rate and listening comprehension. However, although inspection of Figure 7.1 suggests that listening comprehension begins to decline rapidly beyond a rate of 250 wpm, the results displayed in Table 7.3 indicate that losses in listening comprehension do not reach statistical significance until a word rate of 300 wpm is passed. In evaluating the results of significance testing, one must keep in mind the fact that in view of the considerable variance of test scores as indicated by the standard deviations recorded in Table 7.1, relatively large differences among mean test scores would be required for statistical significance. The mean comprehension score of 20.27, obtained at 400 wpm, though quite low, was significantly different from zero, suggesting that there was some comprehension at this word rate. However, in order to be confident that this mean comprehension score had been determined primarily by the listening experience provided the Ss, it would have been necessary to administer the test of comprehension to another group that had not listened to the selection, and this was not done.

The relationship between word rate and listening comprehension, suggested by Figure 7.1 and Table 7.3, is apparently not linear. The hypothesis of linearity was rejected by the test for linearity shown in row 2 of Table 7.2.

Discussion

The results of the present experiment are in close agreement with those of other experiments in which the relationship between word rate and listening comprehension has been studied. In previous investigations (Foulke, et al., 1962; Fairbanks, et al., 1957c), increasing word rate had little effect on listening comprehension below approximately 275 wpm. Increasing word rate beyond 275 wpm resulted in a rapid decline in comprehension. In the present study, the rapid decline in comprehension set in beyond 250 wpm. From a practical point of view,

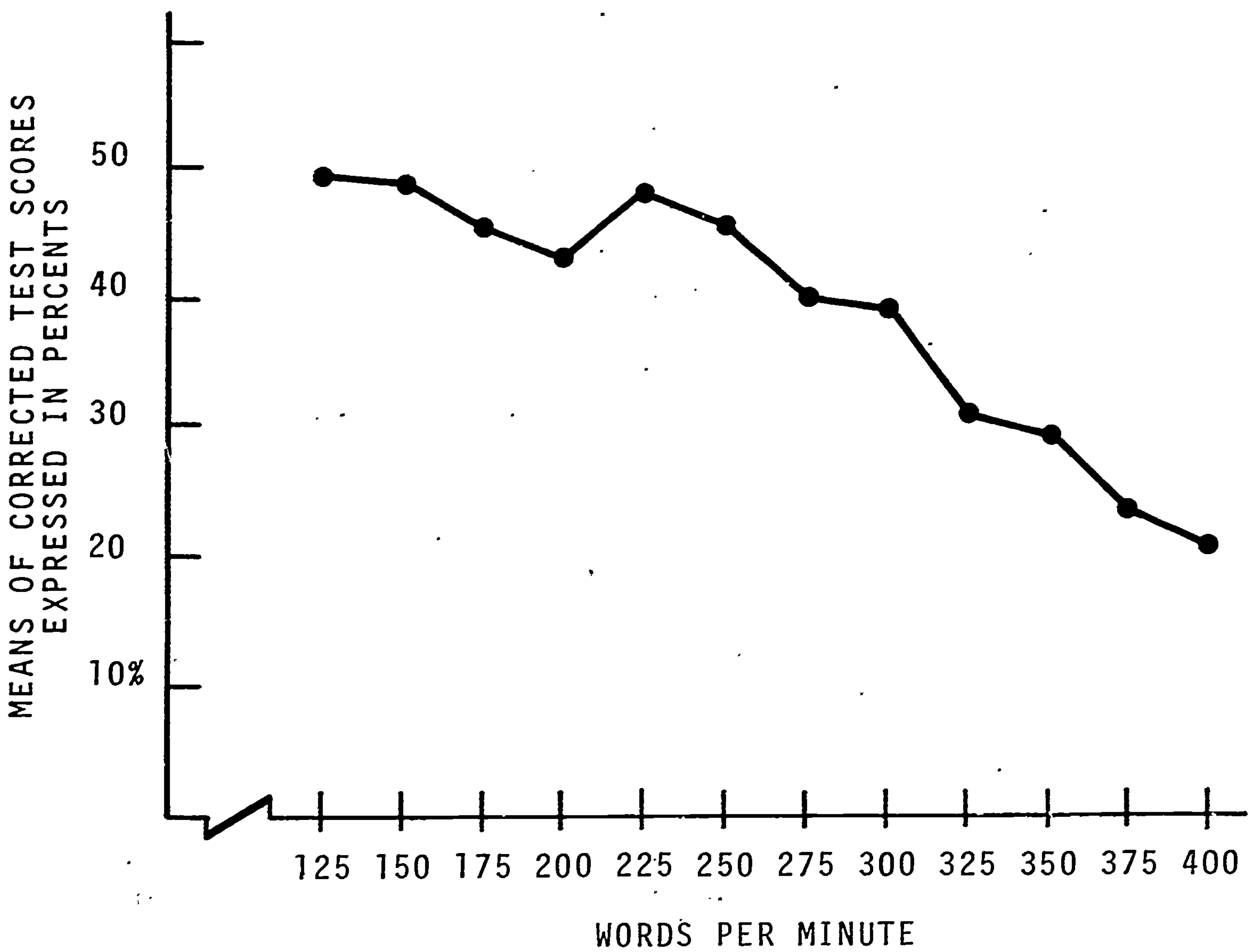

Figure 7.1 Listening Comprehension as a Function of Word Rate

this study, because of the large number of Ss employed, and because of the large number of word rates at which comprehension was determined, provides a firmer basis for making recommendations regarding the accelerated word rates that might safely be considered in those situations in which speech, compressed in time by the sampling method, is to be used to promote faster aural communication. Of course, relevant experience might be expected to bring about some improvement in the ability to comprehend accelerated speech, and the Ss in this experiment had no such experience prior to the experiment. Voor and Miller (1965), for instance, found a slight improvement in comprehension during initial practice trials. The results of other training experiences have been equivocal. Foulke (1964a) found no improvement due to training under any of four conditions of practice. Orr and his co-workers (Orr, et al., 1965; Orr & Friedman, 1967, 1968) have demonstrated significant improvement in the comprehension of speech presented at approximately 425 wpm. However, training experiences have not yet been devised that will result in good enough comprehension of very rapid speech (400 wpm) to permit its practical application in educational settings, and other situations in which people rely on listening. Until successful training methods are developed, the present findings should constitute a fairly accurate picture of the relationship between word rate and listening comprehension.

The present findings also support a hypothesis suggested by Foulke and Sticht (1967) regarding the perceptual problems that accelerated word rates create for the listener. According to this hypothesis, the loss in comprehension that attends an increase in the word rate of speech which has been accelerated by the sampling method, is due not only to a degradation in word intelligibility, but also to a reduction in the perception time needed by the listener to process incoming speech information. Two kinds of evidence can be cited in support of this hypothesis. First, it has been shown (Garvey, 1953b; Fairbanks & Kodman, 1957; Kurtzrock, 1957) that word intelligibility remains at a high level well beyond the compression in time at which the comprehension of connected discourse has begun to decline rapidly. Secondly, the experiments cited earlier in this article in which listening comprehension was determined as a function of word rate, including the present experiment, suggest that listening comprehension is little affected by increasing word rate until a word rate in the neighborhood of 250 or 300 wpm is reached, but substantially affected thereafter. It appears that word rate can be increased, to some extent, without depriving the listener of the perception time required to process speech input. However, beyond a certain point, the available perception time is no longer adequate, and comprehension begins to decline rapidly.

CHAPTER VIII

A SURVEY OF THE ACCEPTABILITY OF RAPID SPEECH*

by

Emerson Foulke

Abstract

In order to gauge the acceptability of time compressed recorded speech for the purpose of reading by listening, a record containing specimens of time compressed speech, and a questionnaire were sent to each of the members of a representative sample of the population consisting of those who use the service offered by Recording for the Blind. Analysis of the responses of those who completed and returned the questionnaires indicated that: a) little practice was required in order to adjust to the task of listening to moderately compressed speech; b) word rates in the neighborhood of 250 or 275 wpm could be understood without difficulty; c) the acceleration of word rate would be more suitable for reading matter that was not of a technical nature; and, d) most readers would listen to books at a faster than normal word rate, if books prepared in this manner were available.

The blind reader is confronted with a serious problem because he must progress at a slow rate. A practiced, adult braille reader can be expected to read at 104 wpm, on the average (Foulke, 1964b). When he listens to material read by a professional reader, he is receiving information at a rate of approximately 175 wpm. On the other hand, many practiced adult readers of print read at a rate of four or five hundred wpm, or even faster.

---

*The material in this chapter also appears as an article in The New Outlook for the Blind, 1966, 60, 261-265.

ERIC

The slow rate at which the blind person must receive written information is more than a nuisance. We live in a highly complex society which is continuously changing and rapidly increasing in complexity. For an individual to react to and participate effectively in this society, he must be informed. He must keep abreast of developments on many fronts, and to do so he must read, and read voluminously.

In addition to these general demands, the individual who must keep informed about developments in a field of knowledge related to his profession or line of work must cope with an increasingly heavy reading burden. There has truly been an information explosion in all fields. The blind person, whose rate of receiving written information is well below 200 wpm, is poorly equipped to deal with his problem. There just is not enough time in the day for him to do the reading he must do to stay afloat. Furthermore, written information is accumulating at a geometric rate, so that his problem becomes progressively more acute.

An obvious solution to this problem is to increase the information transmission rate in whatever communication system the blind person uses. Although research may indicate a way of increasing the braille reading rate, the method for doing this is not now apparent. However, information may be transmitted more rapidly by ear than by touch, and the widespread use of recorded material by blind readers has meant a significant amelioration of their reading problem.

The reading rate of the person who reads by listening has generally been set by the rate at which his oral reader, live or recorded, speaks. There are at least three ways in which this rate might be increased. First, the oral reader could be instructed to read and speak more rapidly. However, when the oral reading rate is increased in this way, the reader soon begins to have difficulty with articulation, phrasing, and inflection. Another method, with which many people have had at least brief experience, is the reproduction of recorded speech at a faster record or tape speed than the speed at which it was recorded originally. By this method, any desired word rate is achieved. Unfortunately, as the word rate increases, there is a progressive distortion in the pitch and quality of the speaker's voice.

The third method is a sampling technique in which parts of a recorded message are reproduced. If the discarded segments of the message are small enough, the human ear cannot detect their absence and the result is accelerated speech that is not distorted with respect to pitch or quality of the speaker's voice. This sampling method may be accomplished manually by cutting out small pieces of the recorded tape, and by joining the cut ends together again. As a matter of fact, the use of periodic sampling to accomplish the time compression of speech was first demonstrated by a splicing procedure. However, cutting and splicing tape is so time consuming that use of such a procedure would invalidate the sampling method for practical purposes. Fortunately, an instrument called the Tempo Regulator accomplishes the time compression of tape recorded speech by periodically failing to reproduce short segments of the recorded tape and by eliminating resulting gaps in the message. This results, like the tape splicing procedure, in speech that is accelerated without distortion in pitch or voice quality. A recorded tape can be reproduced by the Tempo Regulator at any word rate, either slower or faster than the word rate at which the material was recorded. Since the sampling method allowing the time compression of speech produces an output that is relatively undistorted in pitch or voice quality, the result is more pleasing to hear than speech in which time compression has been accomplished by a fast playback speed.

For the past five years, a project has been underway at the University of Louisville to explore the possibility of more rapid aural communications by means of the kind of accelerated speech produced by the Tempo Regulator. In our first study, (Foulke, et al., 1962) we showed that blind school children, in the sixth, seventh, and eighth grades, without prior experience in listening to rapid speech, were able to demonstrate good comprehension of unfamiliar prose presented at a rate of 275 wpm. At higher rates than this, their comprehension began to fall rapidly.

Much of the research since this initial study has been conducted with a view to discovering an effective training procedure that will enable listeners to comprehend very rapid speech (375 wpm or faster). Though we are not able to recommend such a training procedure yet, we have accumulated a good deal of experience in listening to rapid speech and in measuring the comprehension resulting from time compressed speech. One generalization warranted by this experience is that the average listener, without special training, can understand most kinds of reading matter at a rate of approximately 275 wpm. This word rate is a significant improvement over the word rate experienced by the person who reads by listening to conventional, uncompressed recordings. It is a dramatic improvement when compared to the word rate characteristic experienced braille readers. The facts suggest that, without further development, compressed speech can be put to immediate practical use. A reasonable next step would be to present specimens of compressed speech for evaluation by a sample of listeners representative of the people who experience the reading demands which would make compressed speech especially useful. Such an undertaking is reported in the following paragraphs.

Method

Several brief listening selections were chosen. These selections were recorded on magnetic tape by professional readers at the recording studios of the American Printing House for the Blind. The tapes were reproduced on the Tempo Regulator at the desired accelerated word rates and the output of the regulator was recorded on magnetic tape. This "master tape" was then transcribed onto seven-inch vinyl discs by the recording studio of Recording for the Blind. Such discs are used by Recording for the Blind in preparing the recorded texts it distributes to its subscribers. The discs are recorded at 16 2/3 rpm with a playing time of 27 minutes per side. Table 8. 1 describes the contents of the records.

TABLE 8. 1

LISTENING SELECTIONS USED IN SURVEY

Side 1

1. A Hole in the Bottom of the Sea by Willard Bascom; Doubleday; read by Livingston Gilbert; Word Rate, 180 wpm; Listening Time. 40 seconds.

2. A Hole in the Bottom of the Sea by Willard Bascom; Doubleday; read by Livingston Gilbert; Word Rate, 225 wpm; Listening Time, 2 minutes.

3. A Hole in the Bottom of the Sea by Willard Bascom; Doubleday; read by Livingston Gilbert; Word Rate, 275 wpm; Listening Time, 2 minutes.

TABLE 8. 1 (continued)

4. A Hole in the Bottom of the Sea by Willard Bascom; Doubleday; read by Livingston Gilbert; Word Rate, 350 wpm; Listening Time, 2 minutes 7 seconds.

5. Athenian, Spartan and Roman Education by Will Durant from Ideas and Backgrounds; American Book Co.; read by Livingston Gilbert; Word Rate, 275 wpm; Listening Time, 1 minute 4 seconds.

6. Athenian, Spartan and Roman Education by Will Durant from Ideas and Backgrounds; American Book Co.; read by Terry Hayes Sales; Word Rate, 275 wpm; Listening Time, 1 minute 8 seconds.

7. Lost Cities and Vanished Civilizations by Robert Silverburg; The Chilton Co.; read by Livingston Gilbert; Word Rate, gradually accelerated from 180 wpm to 350 wpm; Listening Time, 6 minutes 22 seconds.

Side 2

1. The Battle of New Orleans by Donald Barr Chidsey; Crown Publishers, Inc.; read by Livingston Gilbert; Word Rate, 300 wpm; Listening Time, 20 minutes 30 seconds.

A questionnaire was constructed with questions intended to elicit relevant information about the listener and about his reactions to the compressed listening selections contained on the record. The questionnaire follows.

TABLE 8. 2

QUESTIONNAIRE COMPLETED BY SUBJECTS

Name____________________Date of Birth__________Sex_______
Last year of school or college completed____________________
Degrees received____
Present occupation or profession____________________________

******

TABLE 8.2 (continued)

1. Would you listen to material prepared in this manner if it were available? ___________
2. The material you have listened to has included samples at several word rates. Which word rate did you find most satisfactory? ________ ______________________
3. Judging from samples you have heard, for what kinds of materials do you think the technique of compressed speech would be most suitable? Least suitable? ______________________________ ______________________________ ______________________________
4. You have heard the compressed speech of two different readers. Which reader was most easily understood? ______________________ ______________________________
5. When listening to compressed speech, do you have any preference regarding the sex of the reader? ______________________
6. One of the samples you heard commenced at a normal word rate which was increased slowly to 350 words per minute. What are your reactions to this manner of introducing passages of compressed speech? ______________________________
7. Did you find practice helpful in the understanding of rapid speech? ___________
8. Do you think that you would retain the information presented at fast word rates as well as that presented at a normal word rate? ___________
9. Complete the following (check). I do my reading by means of recordings ______rarely______frequently_______most of the time ______all of the time.
10. Please use the space below for any additional comments that you care to make.

Subjects

A sample of 200 names was drawn from the population of college student subscribers to the service offered by Recording for the Blind. The file from which cards were drawn was organized by states; to insure broad geographic representativeness, one card was drawn at random from each state. This procedure was repeated until the required sample size of 200 was reached. The individuals whose names appeared on these cards were invited, by mail, to participate in the survey. Willingness to participate was indicated by returning the addressed postcard included in the envelope received by the prospective S. Listening samples and questionnaires were sent to the 100 individuals who returned postcards. By completing and returning their questionnaires, fifty-one of these qualified as Ss.

Most of the states were represented in this final sample. The youngest S was 14 years old and the oldest was 56. Thirty-two of the Ss were between 16 and 35 years of age. College students were most numerous but there were some high school students, and four individuals with advanced degrees. Twenty-six of the Ss listed themselves as students, six as teachers, four as members of other professions, two as laborers, one as a business man, and one as a housewife. Eleven Ss did not indicate an occupation or profession.

Procedure

Each person who, by returning his postcard, had indicated a willingness to participate in the survey, was sent an envelope containing the record with samples of compressed speech, plus a braille and a print copy of the questionnaire and instructions for participating in the survey. Participants were given the option of writing their answers to the questionnaire in braille, in print on the appropriate spaces on the braille questionnaire form, or in print in the appropriate spaces on the print questionnaire form. A stamped and addressed envelope was provided for returning the completed questionnaire.

Results and Discussion

For the first question, 92% of the Ss indicated that they would listen to material, the word rate of which was accelerated by the Tempo Regulator method. Answers to question 2 were distributed as follows: 25% of the Ss preferred speech compressed to a rate of only 225 wpm, the smallest amount of compression to which they were exposed. Nevertheless, it was 45 wpm faster than the word rate of the selection before compression. Forty-five percent or nearly half of the Ss judged the rate of 275 wpm to be most satisfactory. This finding is not surprising in view of our previous research in which we found 275 wpm to be the fastest rate at which untrained listeners could demonstrate good comprehension of accelerated speech (Foulke, et al., 1962). Twenty-three percent of the Ss chose 300 wpm as the preferred rate while only 8% favored 350 wpm. This finding is also consistent with the results of the study just cited. The relation reported in this study was one in which comprehension began to fall off rapidly beyond 275 wpm. The responses to questions 1 and 2, considered together, indicate clearly that readers will accept accelerated recordings in which the acceleration, though moderate, is sufficient to accomplish a significant savings in listening time.

An individual's willingness to accept "rapid speech" may depend, in part, upon the amount of reading he must do, and this, in turn, may depend upon his educational level. Therefore, to make an estimate of the influence of educational level upon the willingness to accept "rapid speech," Ss were sorted into two groups according to their educational level. Group 1 consisted of all the Ss who had had one year of college or less, while Group 2 included all of the Ss who had had more than one year. The members of each group were then examined in terms of their responses to questions 1 and 2.

Ninety-six percent of the 26 students with one year of college or less said that they would read material presented at accelerated word rates if it were available. Four percent said that they would not. Eighty-six percent of the group with more than one year said that they would read such materials, and 14% said that they would not. Thus, there is a suggestion that people with more than one year of college are somewhat more reluctant to accept "rapid speech" than those with less education. The difference is small and probably not significant, considering the size of the samples involved, but it deserves further exploration.

TABLE 8.3

LISTENING WORD RATE PREFERENCES AT

TWO EDUCATIONAL LEVELS

Column 1 Word Rate (wpm)	Column 2 One Year of College or Less	Column 3 More Than One Year of College
225	24%	27%
250	12%	0%
275	48%	32%
300	16%	27%
350	0%	14%

The picture is somewhat different when the word rate preferences for these groups are examined. Table 8.3 shows the way in which the two groups distributed their estimates of the most satisfactory word rate. The reader's attention is drawn especially to the difference between the two groups at the faster word rate. It is clear that the group with more education is more willing to accept material prepared at faster word rates. Whether or not this willingness reflects a genuine ability to comprehend material at faster word rates is a matter to be determined by experiment rather than by survey. One explanation of the observed distribution of responses from the two groups to this question may be a keener awareness on the part of the group with more education about the reading problem confronting blind readers.

Because of the wording of question 3, responses to it were varied. However, their general import was clear. Ninety-eight percent of the Ss felt "rapid speech" would be most valuable for narrative and non-technical exposition. Ninety-three percent felt that "rapid speech" would be least suitable for novel and technical information. This finding is also consistent with the results of the study by Foulke, et al., (1962), cited previously, in which the comprehension of a short story presented at accelerated word rates was shown to be better than the comprehension of a scientific selection also presented at accelerated word rates. However, material such as the short story just mentioned is comprehended better by most listeners than scientific information, regardless of word rate. There may be some tendency for a listener, when given the opportunity, to attribute difficulty in comprehension to the manner of the material's presentation.

In the fourth question, 55% of the Ss found the female reader easier to understand, while 45% found the male reader easier to understand. However, the situation is somewhat altered when we consider the responses to question 5. In answering this question, 64% of the Ss expressed a sex preference and, of this group 68% preferred male readers in general while 32% preferred female readers. The finding that male readers are preferred by most listeners is consistent with the experience of those involved in the Talking Book program. Those who listen to Talking Books have, in general, rendered an opinion in favor of male readers. The finding that, in response to question 4, the Ss did not vote in accordance with their general preferences may be due to any of several factors. It may be that differences in the reading styles of the particular readers in question were large enough to override general preferences. It may be that samples produced by the two readers were not recorded on the discs listened to by the Ss with equal fidelity.

One of the samples listened to by the Ss was a five-minute selection that was introduced at a normal word rate. The word rate increased gradually until near the end of the selection when it reached 350 wpm. Seventy-one percent of the Ss indicated by their answers to question 6 that they found this manner of introducing "rapid speech" helpful, 25% found it unneccessary or distracting, and the remaining 4% were undecided.

In response to question 7, of the Ss, 91% found the limited amount of practice afforded by the selections to which they listened helpful in learning to understand "rapid speech" while 9% did not. The report of the Ss on this issue is consistent with other research findings. Voor (1962) and Foulke (1964a) report an initial improvement in the comprehension of "rapid speech" with practice. This practice effect is, however, short-lived and is probably little more than a "warm-up" effect.

In the eighth question, 86% of the Ss felt that they would retain information presented at an accelerated word rate. Fourteen percent felt that they would have difficulty in doing so. The answer to a question of this sort is, of course, decided by experiment and not by the opinions of listeners. However, research reported by Foulke (1964a) and by Enc and Stolurow (1960) indicated that there is no special problem regarding the retention of what is learned when the material to be learned is presented at an accelerated word rate. The opinions of Ss on this issue probably do have some bearing on their willingness to accept "rapid speech".

Answers to question 9 were distributed in the following manner: 10% read by listening to recordings rarely; 23% read by listening to recordings frequently; 50% read by listening to recordings most of the time, while 17% read in this manner exclusively. Though a majority of the Ss answered yes to questions 7 and 8, it is interesting to compare the responses of those who rarely read by listening to recordings with the responses of those who read exclusively by listening to recordings. One hundred percent of the Ss who rarely read by listening to recordings found practice helpful and felt that they would retain information presented at an accelerated word rate. On the other hand, only 75% of those who rely on recordings exclusively for their reading shared this opinion. It appears as if extensive experience with reading by listening introduces a note of caution regarding the improvement that might result from an increase in word rate.

The request in question 10 for additional comments did not elicit any new information. In general, the Ss used the opportunity provided by question 10 to reinforce their responses to other questions in the questionnaire. Most of the Ss expressed their approval of "rapid speech" with certain reservations. A frequent recommendation was that "rapid speech" should be reserved for light, non-technical expositions such as those found in magazines. Of course, a few expressed skepticism regarding its usefulness. A few others expressed unqualified enthusiasm. Several Ss commented on the slight echo effect present in the samples of compressed speech to which they listened. They found this echo mildly disturbing and wondered if it could be eliminated. The echo effect appears to be unavoidable with the equipment currently used for speech compression, and it becomes more pronounced at faster word rates. However, its disturbing influence can be minimized by a proper recording procedure in which careful attention is given to the signal-to-noise ratio.

The findings just reported and their interpretation should be regarded with due caution. No statistical tests were performed to gauge the significance of any of the observed differences which were discussed because, in most instances, the conditions necessary for such tests could not be completely satisfied. Many of the subgroups responsible for the percentages used in comparisons were quite small. Many variables that could influence responses to survey questions such as these were uncontrolled.

CHAPTER IX

LISTENING RATE PREFERENCES OF COLLEGE STUDENTS FOR LITERARY MATERIAL OF MODERATE DIFFICULTY*

by

Emerson Foulke and

Thomas G. Sticht

Abstract

College students naive with respect to accelerated speech determined their preferred listening rate for a simple prose selection by means of the Tempo Regulator, a device that permits continuous variation in word rate without distortion in vocal pitch or quality. The mean preferred listening rate was 207 wpm, a rate well above the speech rates typically reported in the literature. From previous data on blind persons, the authors feel it is likely that with experience in listening to accelerated speech, even faster word rates would be preferred with sighted persons also.

---

*The material in this chapter also appears as an article in The Journal of Auditory Research, 1966, 6, 397-401.

ERIC

By means of special equipment, the listener can now control the word rate of the material to which he listens. With commercially available devices it is possible to reproduce previously tape recorded speech at any desired word rate. Since, with the right equipment, word rate may be varied at will, the question is raised regarding the relation between a listener's word rate preference and his ability to comprehend. There is some reason to suspect that a listener may show better comprehension of material presented at a rate other than his preferred listening rate. Nelson (1948) tested for comprehension of selections presented at 125 - 225 wpm. Although listeners preferred 175 wpm, the data suggested a slight inverse relationship between word rate and listening comprehension. Similarly, investigations of the comprehension of accelerated speech (e.g. Foulke, et al., 1962) have shown an inverse relationship between comprehension and word rate. Yet, in a survey conducted by Foulke (1966c) to determine the listening preferences of blind students who had been provided with a variety of samples of accelerated speech, a speech rate of 275 wpm was most often preferred.

These findings suggest that a listener does not necessarily prefer the word rate that yields the most comprehension. However, to date, the estimates of listener preference regarding word rate have been secondary outcomes of experiments seeking answers to other questions. Because of the desirability of clarifying the relationship between word rate preferences of listeners and listening comprehension, a direct examination of word rate preferences was made.

Method

Subjects

Fifty-eight female and 42 male students in introductory psychology courses served as Ss.

Apparatus

Variation in word rate was accomplished by the use of a Tempo Regulator, a device permitting the time compression or expansion of tape recorded speech without distortion in vocal pitch or quality. This is accomplished by a sampling process in which brief segments of the recorded messages are periodically deleted or repeated. The samples in question are short enough so that, in the case of time compression, their deletion is not detectable by the ear and no entire speech sound is lost.

The Tempo Regulator was modified so that it could be adjusted from 0 to approximately 500 wpm by means of a ten-turn potentiometer. Since this potentiometer must be rotated through 3600° in order to cover the entire range of variation, gradual changes in word rate can be accomplished with ease. The Tempo Regulator was equipped with a tachometer so that, by means of a simple conversion chart, the word rate for any given potentiometer setting could be determined accurately. The output of the Tempo Regulator was amplified by an Eico model HF32 amplifier and fed to S's earphones and E's monitoring speaker.

The listening selection used in the experiment was a story of approximately eighth grade reading level as determined by the Dale-Chall Formula for Readability (Dale & Chall, 1948). It was read orally by a professional reader who announces on radio and television and who is employed in the Talking Book program at the American Printing House for the Blind. His reading was recorded in the Talking Book studios on an Ampex model 300 tape recorder, and the resulting tape was reproduced on the Tempo Regulator during the experiment.

Procedure

A method of limits procedure was used to determine S's preferred listening rate. The Tempo Regulator was first adjusted to produce a word rate well below or well above the range in which listening preferences could be expected to fall. The selection was then presented to S who was instructed to direct E's adjustment of word rate until the word rate at which he preferred to listen was reached. For each S, five ascending trials were alternated with five descending trials, and the starting point for each trial was varied randomly in order to preclude order effects. Subject was seated in an IAC model 400 acoustical chamber and communicated with E by means of an intercom.

Results

The mean word rate for both ascending and descending trials was determined for each S (see Fig. 9.1). The distribution has a mean of 207, a median of 203, and a standard deviation (SD) of 24 wpm.

The mean preferred listening rate was 217 wpm for descending trials and 197 wpm for ascending trials. The difference between these means was significant at the probability level of $p < .01$.

Further analysis indicated a mean preferred listening rate of 212 wpm for males and 204 wpm for females, an insignificant difference.

Discussion

The preferred listening rate of 207 wpm found in this study is more than one SD above 175 wpm, the rate at which the selection was read originally. It is from 1 - 3 SD above the oral reading rates and conversational speech rates that appear in the literature (Bocca & Calearo, 1963; Nichols & Stevens, 1957; Goldstein, 1940).

Furthermore, though the interval of uncertainty (mean word rate for descending trials minus the mean word rate for ascending trials) is fairly large, it also lies well above any of the published word rates for oral reading or conversational speech. The interval of uncertainty found in this experiment covered a range of 15 wpm (212 - 197 wpm). Presumably, the listeners in this experiment would find any word rate in this range equally preferable.

We did not compare, for the same listening selection, preferred rate with the most comprehensible rate. However, Foulke, et al., (1962) indicated that at 207 wpm, there should be a moderate decline in comprehension, at least for those naive with respect to accelerated speech. However, it must be remembered that in all studies exhibiting a difference between the most preferred rate and the most comprehensible rate, Ss have had very little experience. Perhaps with appropriate training, both rates would increase and the gap narrow between them.

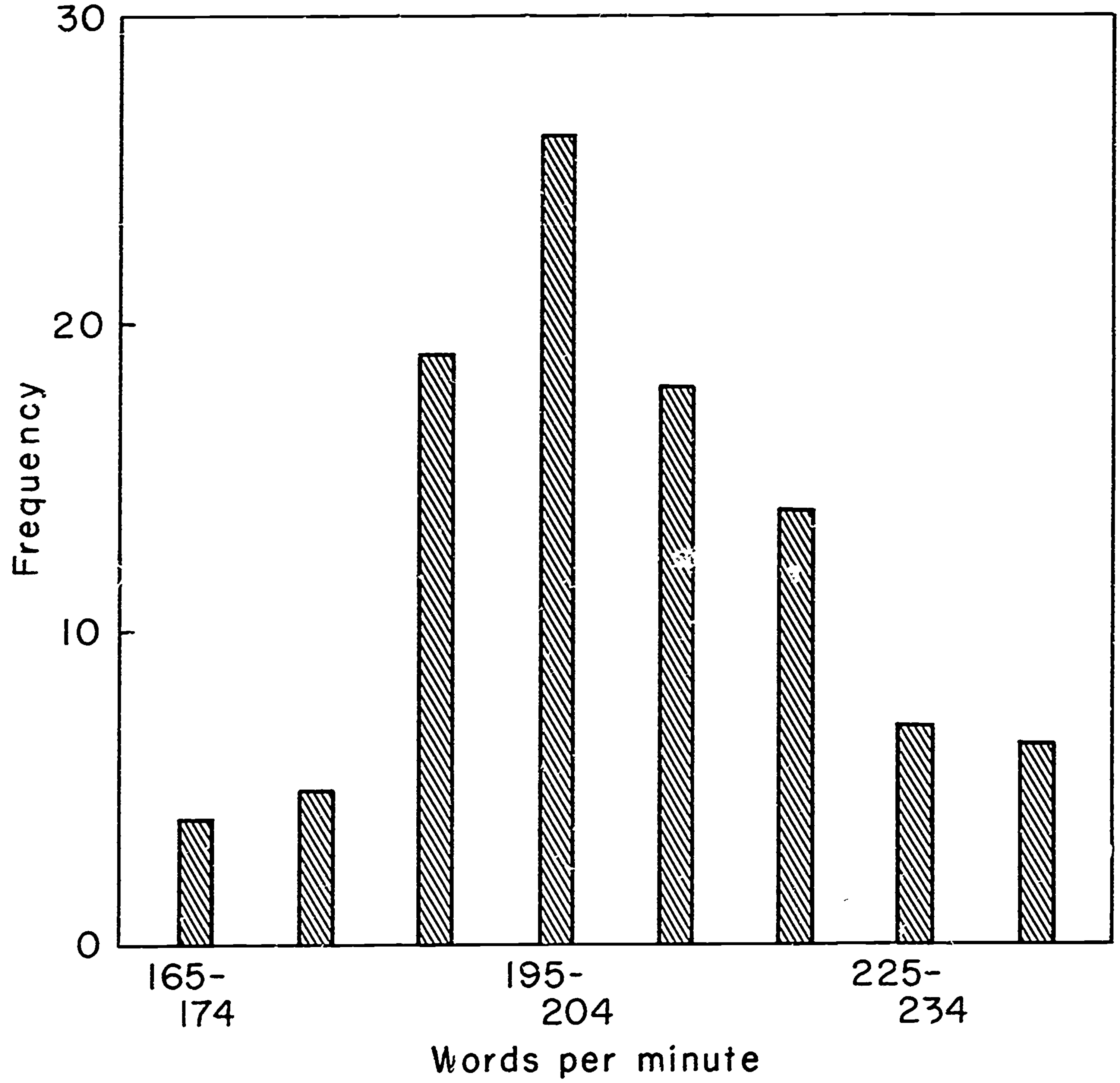

Figure 9.1 Frequency Distribution of Mean Preferred Listening Rates for 100 Ss

The influence that experience in "reading" by listening may have upon the preferred listening rate is suggested by the findings of Iverson (1956) and Foulke (1966c). With a process similar to that used in the present study, Iverson found that many of his 45 blind Ss had difficulty detecting the fact that speech had been compressed by 25% to a word rate of 219 wpm. Most of them estimated that a time compression of 35% to 40% (236 to 245 wpm) was a desirable rate. Foulke found that 45% of a sample of 51 blind Ss judged 275 wpm to be most satisfactory for listening to prose material. It is probable that the faster word rates preferred by blind listeners, as compared to the word rate preferred by the listeners in the present study is due to the fact that blind students must obtain most of their information by listening. Since reading by listening is much slower than silent visual reading, blind students should have more reason to prefer accelerated speech, and more motivation to make effective use of it.

CHAPTER X

THE INFLUENCE OF AGE, GRADE, AND INTELLIGENCE ON THE COMPREHENSION OF TIME COMPRESSED SPEECH

by

Emerson Foulke

Abstract

An experiment was performed to determine the effect of age-grade and IQ on listening comprehension for selections presented at the normal, and two accelerated word rates. The Ss were children drawn from the fifth, eighth, and eleventh grades at residential schools for the blind, and their comprehension was assessed with the STEP Listening Test, the listening passages of which were presented at 175, 275, and 375 wpm. Intelligence was assessed with the WISC, and with the Interim Hayes-Binet Test. The principle result of the experiment was the finding that the maximum word rate at which listening comprehension is preserved depends upon the IQ of the listener. For those Ss in middle and high IQ groups, listening comprehension did not begin to decline seriously until a word rate of 275 wpm had been exceeded. For Ss in the low IQ group, listening comprehension began to decline when the normal word rate of 175 wpm was exceeded.

For blind school children, and others who find it advantageous to read by listening, the ability to compress the time required for the reproduction of recorded oral reading, and hence the ability to increase its word rate, suggests a means of improving this kind of reading. Ordinarily, the reading rate of the person who reads by listening is set by the oral reading rate which is, on the average, 177 wpm (see pg. 106). A person who reads by listening at a rate of 175 wpm, holds an advantage over the typical braille reader, who reads at the rate of 104 wpm (Foulke, 1964b). However, his reading rate does not compare favorably to the silent visual reading rate, estimates of which range between 250 and 300 wpm (Harris, 1947; Taylor, 1937). It should not be the objective of educators to provide for the person who must read by listening an educational experience in which allowance has been made for his slower reading by reducing the reading demands placed upon him. He must be as well prepared by his education for a competitive role in the society at large as his visual reading peers, and to do so, he must have the same opportunity to learn by reading. Yet, the reading demands placed on students in modern educational settings are so heavy that the person who reads by listening finds himself confronted by a shortage of time in which to do the reading expected of him. An obvious solution to this problem is the increase in the word rate of recorded oral reading that is made possible by the techniques of time compression, and research has shown that listeners experience no difficulty in comprehending speech that has been accelerated to 250 or 275 wpm (Fairbanks, et al., 1957c; Foulke, 1968; Reid, 1968). However, this finding is based upon the averaged effects of experimental treatment in experiments in which variables such as educational background, age, and intelligence have either held constant, or were allowed to vary randomly. If those who read, for educational purposes, by listening to time compressed recorded speech are to be school age children, a more satisfactory result will be obtained by taking these variables into account since their effect on behavior is especially pronounced during the developmental years.

Some experiments have been performed in which the effect of word rate on listening comprehension has been determined with age and educational experience serving as parameters. In other cases, although age and educational experience have been held constant in a given experiment, the comparison of experiments in which Ss have differed with respect to age and educational experience may at least suggest the influence of these variables. In those experiments in which school children have served as Ss, age and educational experience have, of course, been varied concomitantly, and the effects of age and educational experience cannot be estimated separately.

Fergen (1954) and Wood (1965) found a positive relationship between the grade level of school children and their comprehension of accelerated speech. Together, their experiments included grades 1, 3, 4, 5, and 6. Since the task of the Ss in Wood's experiment was to carry out the instructions conveyed by short, imperative sentences, one could argue that he was measuring intelligibility, rather than comprehension.

High school and college students have served in many studies in which the influence of word rate upon listening comprehension has been determined (Foulke, et al., 1962; Foulke, 1966b; Foulke, 1968; Fairbanks, et al., 1957a). When the results of these experiments are considered together, there is a suggestion that the relationship between word rate and listening comprehension does not depend very heavily upon age in the age range that encompasses high school and college students. However, because of different experimental materials and conditions, these experiments cannot safely be compared.

The experiments so far reported are not conclusive regarding the effect of intelligence on the comprehension of accelerated speech. Fergen (1954) found no relationship between the IQs of grade school children and the measures of their ability to comprehend accelerated listening selections. However, 230 wpm was the fastest word rate represented in her experiment, and this is a rather moderate acceleration. Wood (1965) found no relationship between IQ and the ability to follow the instructions communicated by short, time compressed imperative statements. However, as previously mentioned, Wood's procedures resemble more closely those used in testing for intelligibility.

There appears to have been no single experiment in which the influence of age, educational experience, and intelligence upon the comprehension of accelerated speech has been assessed. Consequently, an experiment was performed in which blind school children, classified according to age and grade level, and intelligence, were tested for their comprehension of listening selections, presented at several accelerated word rates.

Method

Subjects

Two hundred fifty-six Ss, of both sexes, enrolled in the fifth, eighth, and eleventh grades at eight residential schools for the blind*, served

---

*The writer wishes to thank the superintendents and staff members of the Arkansas School for the Blind, Georgia Academy for the Blind, Illinois Braille and Sight-Saving School, Louisiana State School for the Blind, Maryland School for the Blind, Michigan School for the Blind, Missouri School for the Blind, and Ohio State School for the Blind, for their assistance in the administration of the experiments. The cooperation of the children who served as Ss in the experiment is especially appreciated.

as Ss in the experiment. Although a majority of the Ss were braille readers, some of them were readers of large print. Students who, in the judgment of their teachers, were performing poorly, and whose performance was inconsistent with their grade assignment, were excluded.

Experimental Materials and Apparatus

The tests of listening comprehension used in the experiment were the listening sub-tests of the Sequential Tests of Educational Progress -- Forms 2A, 3A, and 4A. The STEP Listening Test consists of a group of brief listening selections. After hearing each selection, the listener is asked a few questions, of the multiple-choice type. Form 4A is suitable for administration to children in the fourth, fifth, and sixth grades, Form 3A for children in the seventh, eighth, and ninth grades, and Form 2A for children in the tenth, eleventh, and twelfth grades. The listening selections and questions were recorded on magnetic tape by a professional reader in the Talking Book Studios of the American Printing House for the Blind, at 15 ips, by an Ampex tape recorder, model 300. A speech compressor of the Fairbanks type (see pg. 25, ln. 4), constructed at the University of Louisville, was used to alter the word rates of listening selections. When this compressor reproduces tape recorded at 15 ips, it discards periodic samples of the recorded signal that are 40 msec. in duration. The tapes containing the listening selections were reproduced on the speech compressor, as recorded, at 175 wpm (the average oral reading rate), and at 275 and 375 wpm. The output of the speech compressor was recorded on tape at 7 1/2 ips by a Crown tape recorder, model 800. During the experiment, the listening selections in question were reproduced on a Uher tape recorder, model 4000. The output of the tape recorder was distributed to the S's earphones. Subjects listened to experimental materials on Western Electric earphones, type ANB-H-1, which were fitted with circumaural ear cushions to provide isolation from room acoustics. Each headset was provided with a volume control that could be adjusted for a comfortable listening level.

To indicate their answer choices, Ss marked specially prepared braille answer booklets. An entire line was reserved for each test question. A braille number at the left hand margin of each line indicated the question whose answer was to be recorded on that line. Following each number were the letters "A", "B", "C", and "D". To the right of each letter was a small, rectangular enclosure, outlined by braille dots. The pencil marks indicating answer choices were made inside these enclosures. An answer sheet designed in this way enables blind Ss to be certain about the placing of pencil marks. The answer booklets used by readers of large print were replicas of the braille answer booklets, and were printed with stencils typed on a typewriter with bulletin size type.

Procedure

In order to test 256 Ss, it was necessary to visit eight residential schools for the blind. The qualified Ss at each school were distributed throughout all experimental conditions, so that all of the schools were proportionally represented in each experimental condition. In most cases, either Interim Hayes-Binet or WISC IQ scores, obtained some time prior to the experiment by school personnel, were available. These scores were used in the analyses to be reported. The Ss at each grade level were randomly distributed among three experimental groups. Thus, at each of the three grade levels represented in the experiment, there were three comparable groups. The plan was to obtain 30 Ss for each experimental group and this plan was approximately realized. The three word rates at which the STEP Listening Test was presented, were randomly assigned to the three groups at each grade level.

Subjects were tested in classrooms. They were seated at tables and given braille answer booklets or large print answer booklets, and pencils. After they were shown how to adjust their headsets for comfortable wearing, and how to adjust their volume controls for comfortable listening, they heard the recorded instructions for participation in the experiment. The instructions included examples, which provided practice for Ss, and enabled E to assure himself that all Ss understood the test taking procedure. The questions following each listening selection in the STEP tests were read twice. Subjects were told that they could leave questions blank if necessary, but they were advised to guess. Though Ss were told that they could ask for the tape recorder to be stopped after the second reading of a question, in order to allow them more time in which to choose an answer, this request was never made. The tape recorder was stopped occasionally to replace broken or lost pencils, and when this was necessary, the question in progress was completed before it was stopped.

Results

A score, the percent of comprehension test items correctly answered, was determined for each S in the experiment. The means and standard deviations of test scores for the nine treatment groups are shown in Table 10.1. A two-factor analysis of the variance of test scores was performed in order to examine the effects of age-grade and word rate on listening comprehension. The results of this analysis are shown in Table 10.2. The effects of both experimental variables were significant ($p < .01$ in both cases), but their interaction was probably not significant ($p < .10$).

In a second analysis, those Ss for whom IQ scores were available were sorted, at each of the three age-grade levels, into high (110 and higher), middle (90-109), and low (89 and lower) IQ groups. A three-factor analysis of the variance of test scores was then performed with Ss classified according to age-grade level, IQ group, and the word rate of the material to which they had listened. The results of this analysis are shown in Table 10.3. All three experimental variables produced significant effects on listening comprehension ($p < .01$ in all cases). The interaction between IQ group and word rate was significant ($p < .05$). However, the evidence for an interaction between word rate and age-grade level was even less convincing ($p < .25$) than in the first analysis. In order to display graphically the interaction between IQ and word rate, word rate was plotted against mean comprehension test score, with IQ group as the parameter, in Figure 10.1. In this figure, the values scaled on the x-axis are word rates, and the values scaled on the y-axis are test scores, expressed as percents. This figure suggests that: a) in general, comprehension decreased as word rate was increased; b) IQ and listening comprehension were positively related, regardless of word rate; and, c) the word rate beyond which listening comprehension declined rapidly depended upon the IQ of the listener.

Discussion

It is, of course, not surprising to find that children with higher IQs show better listening comprehension. However, the interaction between IQ and word rate is of special interest. For those children in the middle and high IQ groups, listening comprehension was unaffected by increasing the word rate of listening selections from 175 to 275 wpm, but when the word rate was increased to 375 wpm, little or no comprehension was demonstrated. The mean comprehension score obtained at 375 wpm was close to the mean score that would have resulted if Ss had made answer choices at random. This result is consistent with the results usually obtained in experiments in which listening comprehension is measured as a function of word rate (Fairbanks, et al., 1957a;

TABLE 10.1
MEANS AND STANDARD DEVIATIONS OF LISTENING TEST SCORES FOR NINE TREATMENT GROUPS

Grade	Word Rate	Means	SD
	175	64.00	15.57
5	275	68.48	16.18
	375	50.33	11.20
	175	69.03	11.34
8	275	62.50	13.08
	375	54.28	13.17
	175	58.97	12.64
11	275	53.83	14.74
	375	49.09	14.45

TABLE 10.2
ANALYSIS OF VARIANCE OF LISTENING TEST SCORES CLASSIFIED ACCORDING TO AGE-GRADE AND WORD RATE

Source	df	MS	F
A (Word Rate)	2	.46	25.47*
B (Age-Grade)	2	.16	8.71*
AB	4	.04	2.34**
Error	281	.02	

*p < .10
**p < .01

TABLE 10.3
THE ANALYSIS OF VARIANCE OF LISTENING TEST SCORES CLASSIFIED ACCORDING TO WORD RATE, AGE-GRADE AND IQ LEVEL

Source	df	MS	F	p
A (Word Rate)	2	.27	17.90	.01
B (Age-Grade)	2	.17	11.15	.01
C (IQ Level)	2	.45	29.75	.01
AB	4	.03	1.75	.25
AC	4	.04	2.44	.05
BC	4	.01	.70	
ABC	8	.03	1.80	.25
Error	229	.015		

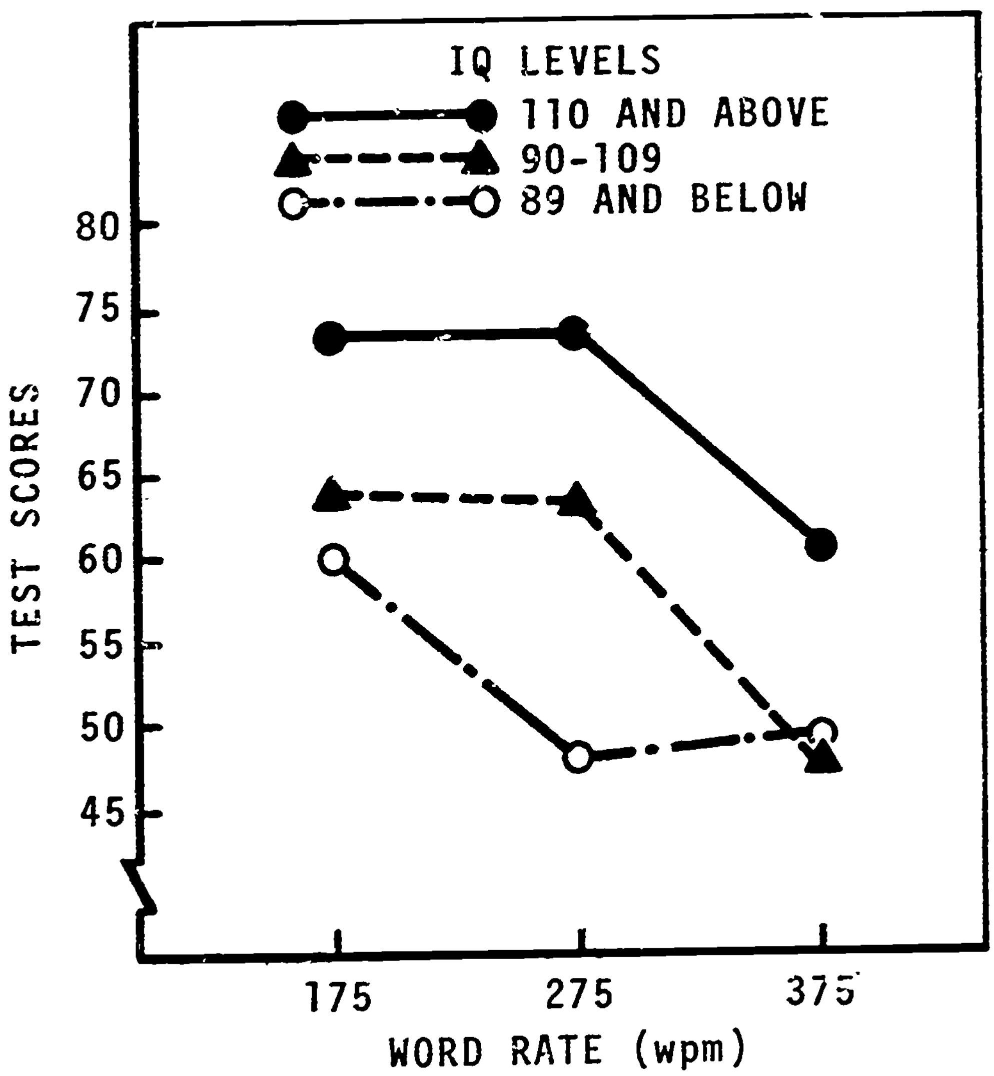

Figure 10.1 Listening Comprehension as a Function of Word Rate at Three IQ Levels

ERIC

Foulke, et al., 1962; Foulke, 1968; Reid, 1968). When the word rate was increased from 175 to 275 wpm, the listening comprehension of the children in the low IQ group declined to a level at or near chance performance, and so was unaffected by a further increase from 275 to 375 wpm. Although the Ss in the low IQ group began to show a loss in comprehension at a lower word rate than the Ss in the middle and high IQ groups, the rate at which comprehension declined with increasing word rate was approximately the same for all three groups. This finding is consistent with the finding reported by Woodcock and Clark (1968). It appears that listeners with low IQs require more time than listeners with high IQs to perform the processing operations mediating the test behavior that is taken as evidence for listening comprehension. The maximum word rate at which there is still enough processing time may depend upon the IQ of the listener, but once that word rate is surpassed, further increases in word rate will cause comprehension to decline at a rate that is similar for all listeners.

To confirm and explicate this apparent relationship, it will be necessary to perform an experiment in which groups of Ss at several IQ levels, which are matched with respect to other important variables, are tested for listening comprehension at several word rates. However, before such an experiment can be performed, there are technical difficulties that must be overcome. An experiment of this sort might require two or three hundred Ss. If these Ss were blind school children, it would be necessary to visit 10 or 15 residential schools for the blind or public school programs in which blind children are enrolled, in order to make up the required compliment of Ss. At present, the only tests available for assessing the intelligence of blind children require individual administration. A testing service is generally available at the schools where blind children are enrolled. However, the attempt to make use of the information about intellectual status provided by this service is frustrating on several counts. The information about intellectual status is obtained by examiners who vary with respect to testing experience, and who may have used different test instruments. The recency of examination varies considerably and, for a variety of reasons, some children have not been examined at all. If the E wishes to have fresh test information in order to conduct an experiment in which IQ is a variable, he must visit the schools in which Ss are enrolled prior to the collection of experimental data, and examine potential Ss individually. Because of the time required for individual examination using test instruments such as the Wechsler Intelligence Scale for Children and the Interim Hayes-Binet Test, if a large number of Ss are required for an experiment, the preparation for such an experiment will be very expensive in time and money.

The solution to this problem is a group test of intelligence, but no such test is available for use with blind children. Consequently, before the experiment outlined above is performed, as well as other similar experiments, an effort will be made to develop a group test of intelligence that is suitable for administration to blind school children. This test will be read aurally by those tested, in order to avoid the difficulties that arise from the considerable variability in braille reading skill which characterizes the population of braille reading children.

The effect of the age-grade variable in this experiment is difficult to interpret. Though significant, the effect was unsystematic. There was a suggestion that Ss in the eleventh grade showed less comprehension than the other Ss in the experiment. However, one would not, on the basis of these results, want to conclude that the ability to comprehend by listening declines with advancing age-grade level. The erratic effect of the age-grade variable may have been the result of uncontrolled differences in the populations sampled at the three age-grade levels, and in the test instruments used at the three age-grade levels. Ideally, a single test instrument should have been used to measure listening comprehension at the three age-grade levels represented in the experiment. However, humans experience considerable development in the age range investigated in this experiment. A single test, suitable for all Ss, would have required items ranging in difficulty from a level suitable for fifth grade Ss to a level suitable for eleventh grade Ss. This is a formidable requirement, indeed. In the present experiment, it was decided to administer to each grade level, the test appropriate for that grade level. It was felt that if the three tests were similar in difficulty, when their listening selections were heard at a normal word rate, the measuring strategy might be adequate to detect an interaction between the word rate variable and the age-grade variable. A reasonable hypothesis might be that, with increasing age and experience, there is a growth in the ability to process the information specified by acoustical stimulation, one consequence of which might be an improvement in the ability to comprehend accelerated speech. If this were the case, it would be expressed as an interaction between the age-grade variable and the word rate variable. Though there was a suggestion of an interaction in the present results, it was not significant, and since the effect of the age-grade variable was not systematic, there is little point in trying to interpret this interaction.

In preparing for the next attempt to investigate the effect of age and grade on listening comprehension at various word rates, an effort will be made to find or develop _ single test of listening comprehension, suitable for administration to Ss in a wide span of school grades; for instance, grades three through twelve. With a test of this sort, and a group aural intelligence test, an experiment similar in plan to the one reported here should yield much more conclusive results.

ERIC

CHAPTER XI

THE ORAL READING RATE

by

Emerson Foulke

Abstract

Two investigations of the oral reading rate were conducted. An oral reading rate of approximately 177 words per minute or 254 syllables per minute was found. The average number of syllables read per minute appeared to be a more stable indication of the oral reading rate than the average number of words read per minute. The oral reading rate was shown to depend upon the book being read, but those differences among books that were responsible for the observed differences in reading rate were not specified. Due to inadequacies in the design of the two studies, the influence of variables pertaining to the oral reader could not be properly assessed.

The rate of occurrence of words in spoken language depends upon both personal and situational factors, and varies widely. Nichols and Stevens (1957) found a conversational speaking rate of 125 wpm. The oral reading rate is usually much faster. Johnson, et al., (1963), found an average oral reading rate of 176 wpm. The oral reading rate is quite variable, and depends upon such factors as the skill of the oral reader and the difficulty of the material he is reading. The oral reading rate is usually the speech rate of interest to those concerned with time compressed speech since in most cases, it is recorded oral reading that is compressed in time.

When those experiments are compared in which listening comprehension has been measured as a function of the amount of compression in time (for example, Fairbanks, et al., 1957c; Foulke, et al., 1962; Foulke, 1968; Reid, 1968), it appears that although the initial or uncompressed word rates of the listening selections used in these studies varied considerably, listening comprehension begins to decline rapidly beyond a word rate of approximately 275 wpm. Since these listening selections were originally read at different rates, different amounts of compression were required to achieve the word rate at which listening comprehension began to decline rapidly.

To confirm this impression, Foulke (1967) performed an experiment in which a listening selection was recorded on tape at three different word rates by a professional reader. The three tapes were then compressed to a final word rate of 275 wpm. There were no significant differences in the comprehension test scores of three comparable groups of Ss who listened to the three compressed renditions of the listening selection.

Evidence of the sort just presented suggests that listening comprehension varies directly as a function of word rate, and only indirectly as a function of the amount of compression in time. It follows that a decision regarding the amount of compression to which a given recorded listening selection should be subjected will depend upon the rate at which it was read originally. In making such decisions, the usual practice has been to assume that the rate at which a listening selection was read probably did not depart significantly from the average oral reading rate of approximately 175 wpm, and to use this value in computing the amount by which a listening selection is to be compressed. In order to justify an assumption of this sort, it is necessary to know not only the average oral reading rate, but also the variability in the measures that determine this average. Furthermore, the contribution to this variability of such factors as the fluctuation in the individual oral reader's speaking rate from time to time, interpersonal differences in reading ability, and differences associated with the material to be read, should be assessed. Accordingly, an investigation was undertaken in order to obtain the information needed for a better description of oral reading behavior.

Study One

In the first of two studies, samples of oral reading were obtained from two sources -- the Talking Book Records distributed by the Library of Congress, and radio newscasts. Each sample consisted of one minute of uninterrupted oral reading. Samples that included unusually long pauses, of the sort that might be introduced by an oral reader between chapters in a book, or between items in a newscast, were not used. The number of samples obtained for each reader varied, depending upon the material available at the time the samples were collected. The results of the survey, for both Talking Book readers and radio newscasters, are shown in Table 11. 1. Readers are designated by their initials, and these initials are entered in column 1. The number of words read during one minute samples are entered in succeeding columns. Mean reading rates are entered in the final column, at the right hand margin of the table. Reading across a row in this table, one first encounters the initials that designate a particular reader, then the number of words he read during each of the one minute samples of his reading that were obtained, and finally, his mean reading rate.

The mean values for oral reading rates shown in Table 11. 1 are in close agreement with the mean values reported elsewhere in the literature. However, there is considerable variability in the word counts upon which these mean values are based. There is often wide variation in the samples obtained from a single reader. There are also apparent differences among readers with respect to word rate. However, such differences might also be due to the kind of material read, and the reading samples in this study were not chosen in such a way that variation due to characteristic differences among readers could be distinguished from variation due to the nature of the material read.

Study Two

In a second study, a more thorough examination of the oral reading rates of Talking Book readers was conducted*. The Talking Books examined in the study were chosen in such a way that the effects of several factors on the oral reading rate might be estimated. Books were chosen from several of the categories of reading matter that are distinguished in the

---

*The author wishes to thank Miss Helen Cannon, the chief librarian at the Wolfner Memorial Library for the Blind, 3844 Olive Street, St. Louis, Missouri 63108, and her staff, for their assistance in obtaining the materials for this study. The Talking Books that were examined for oral reading are in the collection at the Wolfner Library.

TABLE 11.1

A SURVEY OF ORAL READING RATES

Readers	Talking Book Readers Number of Words Read During One Minute Samples										Mean WPM
L. G.	152	161	169								161
A. M.	168	170									169
D. M.	185	200	217								201
A. H.	153	155	184	191	192	193	206	210			186
A. C.	157	161	161	162	172						163
T. C.	165	166	167	167	186						170
R. H.	161	169	174	176	178	182					173
B. B.	154	167	173	174	174	175	188	196	207	224	184
R. D.	139	142	154	163	163						152
R. B.	142	181	182	190	195						178
G. S.	106	144	144	150	156						140
J. C.	169	174	174	183	188						178
W. G.	174	174	186	186	205						185
M. H.	151	151	152	170	173	186					163
A. S.	155	155	159	178	201						169
G. R.	174	195	217	226	227						208
G. W.	159	160	169	184	193						172

Number of Samples = 88
Mean of Samples = 174 wpm
Standard Deviation of Samples = 23.53

Readers	Radio Newscasters Number of Words Read During One Minute Samples										Mean WPM
J. S.	164	174	180								173
W. H.	159	177	180								172
F. L.	157	164	168	175							166
L. T.	149	159	161	166	166	170	173	176	178	179	168
B. W.	158	163	165	168	174	179	192	208			176
B. R.	158	186	191	201							184
N. B.	164	165	169	182	184	187					175

Number of Samples = 38
Mean of Samples = 174 wpm
Standard Deviation of Samples = 13.10
TOTAL NUMBER OF SAMPLES = 126
MEAN OF ALL SAMPLES = 174
STANDARD DEVIATION OF ALL SAMPLES = 17.94

classificatory system used by the Library of Congress. This was done because of the possibility that books in different categories might be different with respect to factors that might affect the oral reading rate, such as vocabulary and syntactic complexity. Samples of oral reading were obtained from many different readers, in order to gain an impression of the variability in reading rate associated with individual differences among readers. In order to gain an impression of the variability in the reading rate of a single reader, samples were taken from five different books, in five different reading categories, read by the same reader.

Method

An investigator was sent to the Wolfner Library. In consultation with the library staff, she identified popular categories of reading matter, and frequently requested Talking Books in each category. Ten one-minute samples of each Talking Book were taken. These samples were distributed more or less evenly throughout the book. Samples were excluded that contained pauses of the sort that might be introduced by an oral reader to indicate boundaries between chapters or other divisions of a book, so that each sample contained continuous speech. The Talking Book records containing the desired samples were reproduced on a record player, connected to a tape recorder, and the samples chosen for use in the study were copied. The tape record produced in this manner was subsequently examined, and the following results were observed.

Results

For each one-minute sample of oral reading, both the number of words and the number of syllables were counted. Table 11.2 shows the means and standard deviations of oral reading rates, in words per minute and syllables per minute, for each book within a category of reading matter, for each category of reading matter, and for all of the samples that were examined. In this table, each of the books from which samples were drawn is designated by a number. Since the reader may wish to judge for himself the extent to which books were representative of the categories from which they were drawn, the titles of all of the books from which samples were drawn are shown in the Appendix. Each book listed in this Appendix is designated by the same number used in Table 11.2. All readers are identified by their initials in this table, and by their full names in the Appendix.

TABLE 11.2

MEANS AND STANDARD DEVIATIONS OF WORD AND SYLLABLE RATES

	Books	Oral Readers	Means		Standard Deviations	
			Words per minute	Syllables per minute	Words per minute	Syllables per minute
FICTION	3	J. W.	215	258	11.05	12.96
	9	H. S.	176	240	12.87	8.17
	7	K. M.	174	253	11.36	21.33
	4	L. G.	193	268	10.94	14.94
ALL FICTION SAMPLES			189	255	20.07	17.68
HISTORY	13	N. R.	206	286	14.26	22.22
	8	S. N.	154	250	10.19	15.78
	2	K. M.	170	258	9.35	15.57
ALL HISTORY SAMPLES			177	264	24.68	23.40
LITERATURE	14	K. M.	178	243	8.09	10.31
	6	W. G.	183	245	13.31	16.05
ALL LITERATURE SAMPLES			180	244	10.92	13.15

TABLE 11.2 (continued)

	Books	Oral Readers	Means: Words per minute	Means: Syllables per minute	Standard Deviations: Words per minute	Standard Deviations: Syllables per minute
PSYCH & PHIL	12	P. C.	193	266	13.10	25.06
	11	N. L.	171	244	10.45	19.65
	5	K. M.	160	272	9.04	9.19
ALL PSYCH. & PHIL. SAMPLES			175	261	17.58	22.07
RELIGION	10	E. R.	173	244	15.52	29.51
	15	K. M.	172	272	13.11	17.51
	1	O. B.	182	242	7.82	18.36
ALL RELIGION SAMPLES			176	253	13.07	25.68
TOTAL SAMPLES			179	254	21.03	36.41

Inspection of the reading rates recorded in Table 11.2 suggests considerable variability in the rates at which different books were read. Analyses of the variance of word rates and syllable rates were perfromed in order to confirm this suggestion, with observations classified according to the books from which samples were drawn. The results of these analyses are shown in Tables 11 3 and 11.4. The books were read at significantly different rates, in terms of both words and syllables per minute ($p < .01$ in both cases). In order to identify the factors responsible for this variability, several additional analyses were performed.

TABLE 11.3

ANALYSIS OF VARIANCE OF WORD RATE WITH

OBSERVATIONS CLASSIFIED ACCORDING TO

BOOKS FROM WHICH SAMPLES

WERE DRAWN

Source	df	MS	F
Between Books	14	2618.82	19.55*
Within Books	135	133.93	

*$p < .01$

TABLE 11.4

ANALYSIS OF VARIANCE OF SYLLABLE RATE WITH

OBSERVATIONS CLASSIFIED ACCORDING TO

BOOKS FROM WHICH SAMPLES

WERE DRAWN

Source	df	MS	F
Between Books	14	1912.34	5.88*
Within Books	135	325.03	

*$p < .01$

Analyses of the variance of word rates and of syllable rates were performed, with observations classified according to the categories of reading matter from which they were obtained. The results of these analyses are shown in Tables 11.5 and 11.6. There were no significant differences in word rate that could be related to the categories of reading matter from which samples were drawn, but differences in syllable rate were significant ($p < .01$). The failure to find agreement between the two indices of reading rate, in this regard, is puzzling. If word length, as measured by average number of syllables per word, varied significantly from category, and if readers produce the syllables in longer words at a faster rate, such a result might be obtained. However, this possibility is not born out by subsequent analysis (see pg. 118, ln. 22).

TABLE 11.5

ANALYSIS OF VARIANCE OF WORD RATES WITH OBSERVATIONS CLASSIFIED BY READING CATEGORIES

Source	df	MS	F
Reading Categories	4	1134.80	2.88*
Error	95	393.95	

*Not significant.

TABLE 11.6

ANALYSIS OF VARIANCE OF SYLLABLE RATES WITH OBSERVATIONS CLASSIFIED BY READING CATEGORIES

Source	df	MS	F
Reading Categories	4	1759.58	3.65*
Error	95	481.61	

*$p < .01$

The oral reading rate should vary, to some extent, as a function of factors pertaining to the oral reader, himself, such as his interpretative style and his background of experience. To test for a relationship of this sort in the data of the present study, analyses of the variance of word rates and of syllable rates were performed, with observations classified according to those oral readers whose productions were examined in the study. The results of these analyses are shown in Tables 11.7 and 11.8. The variations in oral reading

TABLE 11.7

ANALYSIS OF WORD RATES WITH OBSERVATIONS CLASSIFIED ACCORDING TO ORAL READERS

Source	df	MS	F
Oral Readers	9	3189.26	2.53*
Error	90	1258.45	

*p<.05

TABLE 11.8

ANALYSIS OF VARIANCE OF SYLLABLE RATES WITH OBSERVATIONS CLASSIFIED ACCORDING TO ORAL READERS

Source	df	MS	F
Oral Readers	9	2185.00	5.93*
Error	90	368.06	

*p<.01

rate that could be related to differences among oral readers were significant in terms of both word rate (p<.05) and syllable rate (p<.01).

In order to discover whether or not the same oral reader reads different books at significantly different rates, the ten observations in each of the five books read by Kermit Murdock (see books 2, 5, 7, 14, and 15 in Table 11.2 and the Appendix) were examined. Analyses of the variance of word rates and syllable rates were performed, with observations classified according to the five books read by Murdock.

The results of these analyses are shown in Tables 11.9 and 11.10. The effect due to books was significant for both word and syllable rates ($p < .01$ in both cases).

TABLE 11.9

THE ANALYSIS OF VARIANCE OF WORD RATES WITH OBSERVATIONS CLASSIFIED ACCORDING TO DIFFERENT BOOKS READ BY THE SAME ORAL READER

Source	df	MS	F
Books	4	439.62	4.10*
Error	45	107.11	

*$p < .01$

TABLE 11.10

THE ANALYSIS OF VARIANCE OF SYLLABLE RATES WITH OBSERVATIONS CLASSIFIED ACCORDING TO DIFFERENT BOOKS READ BY THE SAME ORAL READER

Source	df	MS	F
Books	4	1532.12	4.08*
Error	45	375.16	

*$p < .01$

Discussion

The data analyzed in this study were obtained from existing specimens of oral reading. Working within this constraint resulted in serious departures from sound experimental design. This study is not as conclusive as it might have been, if recourse to the logic of a well designed experiment had been possible. However, in order to perform such an experiment, it would have been necessary to examine the oral reading of a number of different readers, all of whom read the same books, and it would have been necessary to select these books from categories of reading matter known to differ by known amounts with respect to factors such as vocabulary and syntactic complexity, whose influence on the oral reading rate might reasonably by hypothesized. These conditions cannot be realized by sampling the existing Talking Book literature. Since Talking Books are expensive to produce, books of restricted interest cannot be considered, and books with a broad general appeal are very likely to be similar with respect to vocabulary and syntactic complexity, regardless of the category of reading matter to which they may have been assigned by a librarian. Of course, one would not find, in the Talking Book literature, the same book read by several different readers. In planning an experiment that met the required conditions, one would have to consider the possibility that the experiment might be too expensive in terms of the value of the information it would yield. In the present case, a decision was made to determine what could be learned by examining those specimens of oral reading already available in the Talking Book literature, in spite of the fact that it was usually not possible to choose samples in a manner that permitted independent variation of those factors believed to influence the oral reading rate.

Initially, it was hoped that the system of classification used by the Library of Congress would result in categories of reading matter that were different with respect to such factors as vocabulary and syntactic complexity, so that the effects of these factors on the oral reading rate might be observed. However, since those books chosen for presentation as Talking Books must be generally appealing to a lay reading public, they must contain words and syntactic forms that will be generally understood. Inspection of the books selected for examination in this study revealed no apparent differences in vocabulary and syntactic complexity and, as was shown in the Results section, (see Tables 11.5 & 11.6), no effects due to categories of reading matter were manifested in the results. It can be said that, by examining books drawn from several different categories of reading matter, a population of books of general interest was broadly sampled.

Because oral readers and the books read by them could not be varied independently, it is not possible to extricate their effects on the results of the present study. However, when considered together, the results of the several analyses that were performed in an attempt to identify sources of variation are suggestive. In the analyses in which a single reader's renditions of five different books were examined (see Tables 11.9 & 11.10), a significant overall effect due to differences among books was found. Since the analyses reported in Tables 11.5 and 11.6 indicated that the categories of reading matter from which books were chosen did not have a significant effect on the oral reading rate, the observed differences in oral reading rate must have been due to differences among books within categories. Books might differ in a variety of ways, but the relevant differences in this case would probably relate to such factors as vocabulary and syntactic complexity. This conclusion must be tempered by the possibility, not demonstrated in the results of the present experiment, of an interaction between reader variables and reading matter variables. It might be, for instance, that a reader with a larger vocabulary and more experience with complex syntax could read a selection with complex syntax and with a relatively large number of long and infrequently occurring words more fluently than another reader without his background, whereas the two readers might read a selection with simple syntax and limited vocabulary equally well. Furthermore, two oral readers, equally skilled with regard to these factors, might, because of different interpretative styles, read the same book at different rates. However, in the present study, only professional readers, with years of oral reading experience, were used. Considering the books they read, it is unlikely that any of these readers were embarrassed by unfamiliar vocabulary or syntactic complexity. Furthermore, many of the readers who produced the samples of oral reading examined in this study are radio and television announcers, and their interpretative styles are similar.

If reading matter variables, such as vocabulary and syntactic complexity, were responsible for the differences in the results analyzed in Tables 11.9 and 11.10, there is no reason to believe that they did not also contribute to the results analyzed in Tables 11.7 and 11.8 where, although observations were classified according to oral readers, each oral reader read a different book. In fact, since all of the readers were professionally trained, and since many of them had similar professional backgrounds, reading matter variables may have been primarily responsible for the significant differences revealed by these analyses as well. To pursue this question further, it would be necessary to arrange for the same reading selections to be read by different oral readers.

The reading matter variable of vocabulary, already mentioned, can be further analyzed into component variables, such as frequency of word usage, phonetic structure, and word length. Words that occur more frequently in general English usage may be more familiar to the typical oral reader who may, as a result, identify and pronounce them more rapidly. Words with different phonetic structures may place different, and more or less strenuous articulatory demands upon the oral reader, who may be able to render some phonetic structures more facilely and rapidly than others.

Word length is a variable of particular interest because of its implications for the measure used in assessing reading rate. If oral readers produce speech sounds at a fairly constant rate, two reading selections, differing in average number of syllables per word, should be read at different word rates, but similar syllable rates. Since reading matter does vary from selection to selection with respect to number of words per syllable, the average number of syllables read per minute might, as Carroll's data suggest (Carroll, 1967), be a more stable indication of the oral reading rate than the average number of words read per minute. If this is the case, it should be reflected in the present results.

If oral readers produce syllables at a fairly constant rate, regardless of the average number of syllables per word, increasing the average number of syllables per word should result in a decrease in the oral reading rate, when it is expressed in words read per minute, but not when it is expressed in syllables read per minute. To examine this proposition, the mean syllable values in Table 11.2 were divided by the mean word values recorded in the same table to obtain the average number of syllables per word for each of the books from which samples were drawn. This information is presented in Table 11.11.

The correlation between average number of syllables per word (see Column 3, Table 11.11) and the average number of words read per minute (see Column 2, Table 11.11) was assessed by the Pearson Product-Moment formula, and an $\underline{r}$ of minus .78 was found. This is a fairly strong degree of relationship, and it indicates, as expected, that as the average number of syllables per word is increased, the average number of words read per minute decreases. An $\underline{r}$ of .27 was found when average number of syllables per word (Column 3, Table 11.11) was correlated with average number of syllables read per minute (Column 1, Table 11.11). An $\underline{r}$ of this magnitude is not significantly different from zero with a sample size of 15. This lack of

TABLE 11.11

AVERAGE NUMBER OF SYLLABLES PER WORD

FOR 15 DIFFERENT BOOKS

Books	SPM*	WPM**	SPW***
3	259	215	1.20
9	240	176	1.36
7	253	174	1.45
4	268	193	1.39
13	243	178	1.37
8	245	183	1.34
2	244	173	1.41
14	273	172	1.59
6	242	182	1.33
12	266	193	1.38
11	244	171	1.43
5	272	160	1.70
10	286	206	1.39
15	250	154	1.62
1	258	170	1.52

* SPM = Syllables per minute
** WPM = Words per minute
***SPW = Syllables per word = SPM/WPM

relationship indicates that, as the average number of syllables per word is varied, oral readers produce syllables at a more constant rate than words.

One consequence of the fact that syllables are read at a more constant rate than words should be a smaller coefficient of variation ($V = (\sigma / M) \times 100$) for the distribution of observations of the number of syllables read per minute than for the corresponding distribution of observations of the number of words read per minute. The coefficient of variation for the 150 observations of words read per minute (10 samples from each of 15 books) was 11%, and it was 9% for the corresponding distribution of observations of syllables read per minute. Thus, although the difference between the two coefficients of variation was small and possibly not significant, it was in the expected direction, and it suggests that syllables are produced at a more constant rate by oral readers than words.

Conclusions

Several conclusions appear to be warranted by the results of Studies One and Two. The average oral reading rate for skilled oral readers, when assessed in terms of the number of words read per minute, is approximately 177 wpm. There is considerable variability in the number of words read per minute by different readers or by the same reader reading different books. Oral readers produce syllables at a more constant rate than words. If careful specification of the oral reading rate is required, this specification should be made in terms of the number of syllables read per minute, in preference to the number of words read per minute. Different books, presumably differing with respect to such factors as vocabulary and syntactic complexity, are read at different rates. In order to specify further the contributions of these and other reading matter variables, it will be necessary to perform research in which the reading passages read by oral readers are quantitatively different in known ways. The results of the two studies reported in this chapter do not permit definite conclusions regarding the effects of variables pertaining to the oral reader. To assess the effects of these variables, it will be necessary to perform studies in which different oral readers render the same reading matter.

ERIC

CHAPTER XII

OTHER EXPERIMENTS

by

Emerson Foulke

Abstract

In the course of this project, several experiments were undertaken that are reported only in summary form. In some cases, the experiments were too minor in scope to merit a full, detailed report. In other cases, because of problems due to insufficient staff, equipment inadequacy, or unavailability of Ss, experiments were interrupted before completion. In still other cases, experiments were not completed at the writing of this report. These experiments included: Compressed Speech Viewed as a New Language; Separating the Effects on the Comprehension of Accelerated Speech of Decreasing Word Intelligibility and Increasing Word Rate; Effects of Stimulus and Interstimulus Duration on the Immediate Recall of Time Compressed Sequences of Different Orders of Approximation to English; Forward Versus Backward Reproduction of Tapes Compressed by the Electro-mechanical Sampling Method; and, The Experimental Control of Listening Difficulty.

During the course of this project, several experiments were initiated that will not be reported in detail. In some cases, although data collection was completed, it proved impossible to complete data analysis and to prepare a detailed account in time for its inclusion in this final report. In other cases, experiments were discontinued because preliminary findings did not seem promising, because of technological problems, or because Ss were unavailable.

Completed Experiments

Compressed Speech Viewed as a New Language

As speech is compressed in time, and its word rate is accelerated, a point is reached beyond which it is no longer comprehensible to a listener. Of course, practical benefits of considerable importance would be realized if listeners could be taught to understand speech presented at an incomprehensibly fast rate. Several investigators (Foulke, 1964a; Voor & Miller, 1965; Orr, et al., 1965) have evaluated training experiences designed to improve the comprehension of accelerated speech. These experiences have consisted of little more than simple exposure, and their success has not been remarkable. This limited success may be the consequence of an upper limit on the rate at which the listener can process speech. Rates that exceed this limit may simply exceed his perceptual capacity. On the other hand, the training experiences so far evaluated may have been too ingenuous in their conception. If listeners are to be taught to comprehend accelerated speech, it may be necessary to analyze the task of comprehending such speech into its component skills, and to formulate training experiences which promote acquisition of these skills.

Discrimination is prerequisite to the comprehension of normal speech. In order for a listener to identify words, he must be able to discriminate one word from another. As words are compressed in time, the resemblance between them and their uncompressed counterparts is decreased. A point is reached beyond which they are no longer identifiable. Further more, the listener cannot discriminate among them, except in a gross sense. He may be able, on the basis of duration alone, to distinguish between a one-syllable and a two-syllable word, but he cannot tell two one-syllable words or two two-syllable words apart. However, practice in listening to the unfamiliar sounds resulting from the compression of speech, under appropriate conditions, may restore his lost ability to discriminate and identify. The listener may be able to comprehend time compressed speech composed of time compressed words and phrases he has learned to identify.

To explore this possibility, a few Ss were given practice in the identification of highly compressed, common words. Several 50-member groups of words were drawn from the 1,000 most frequently occurring words in the Thorndike-Lorge Count (1944). Each group contained a mixture of nouns, pronouns, adjectives, verbs, and adverbs. These words were recorded on tape and compressed to 35% of their original durations. Subjects learned to identify the 50 words in each group by a paired-associates procedure. A trial consisted of one presentation, in random order, of the 50 words in a group. Each S attempted to identify each word. If his guess was correct, he was so informed. If it was incorrect, he was informed of the correct response. According to plan, each S was to receive practice on a particular group of words until he reached a criterion of two successive errorless trials. Occasionally, however, after many trials, an S appeared to be unable completely to eliminate errors. To prevent his discouragement, he was advanced to the next stage of practice without having met the criterion of mastery. When a group of words had been learned, S was given practice in identifying simple, time compressed sentences formed from the words in the group. Following this, he was introduced to a new group of words. When this group was mastered, he was given practice in identifying time compressed sentences formed from the words in this, and all previously mastered groups.

The results of this experiment have not yet been completely analyzed, and it is only possible to report those obvious impressions gained by inspection of the data. In spite of the fact that word intelligibility was assured by the practice given Ss, they showed very poor understanding of the sentences composed with these words. The poor performance on sentences was quite resistant to practice. Even after many trials with sentences, Ss were unable to understand as many as half of them. Although Ss could recognize many of the words in sentences, the order in which words were recalled was frequently incorrect. These findings suggest that when speech rate is too high, the demands upon a listener's ability to perform those processing operations involved in the understanding of spoken language may be excessive. The operations involved in the perception of spoken language require time, and if not enough time is allowed for these operations, comprehension will deteriorate.

If practice enables a listener to identify highly compressed words, practice of the right sort may also enable him to increase the rate at which he can perform the processing operations involved in the comprehension of accelerated speech. In an experiment suggested by the outcome of this experiment, Ss will learn to identify compressed words, presented in isolation. However, when sentences are composed with these words, each word will be separated from its neighbors by unfilled time intervals. As practice in identifying these sentences continues, the intervals between words will be shortened gradually, until a time compressed version, without added time, is reached. It is hoped that a practice schedule of this sort will enable listeners to perform those processing operations involved in the comprehension of accelerated speech at a faster rate.

Separating the Effects on the Comprehension of Accelerated Speech of Decreasing Word Intelligibility and Increasing Word Rate

As speech is compressed in time, there is a loss in listening comprehension. This loss is probably due, in part, to a decline in the legibility of the speech signal and, in part, to an increase in the rate of occurrence of speech signals. In an effort to separate these effects, Miss Ruth Ann Overmann performed an experiment, to be reported in her master's thesis, in which the word rate of several compressed listening selections was varied by varying the amount of pause time at phrase and sentence boundaries. The selections contained in the Nelson-Denny Tests of Reading Comprehension were recorded on tape and compressed to three different fractions of original production time. At each compression, two test tapes were prepared. The word rate of one of each pair of test tapes was restored to the original or uncompressed word rate by inserting pause time at phrase and sentence boundaries. The other member of each pair was simply the compressed version, with no pause time added. Thus, at each compression represented in the experiment, the two versions of the listening selection were alike with respect to the magnitude of compression of individual words, but unalike with respect to pause time and word rate. If listeners use the pause time distributed throughout fluent speech to perform needed processing operations, one would expect those listeners who heard the selections with pause time added to show better comprehension than the listeners who heard the compressed selections in which no pause time had been inserted. The results of the experiment were in general agreement with this expectation. In no case did the group of Ss who listened to compressed tapes with pause time added comprehend the selection as well as a control group who heard the uncompressed version of the listening selections. However, in every instance, the insertion of pause time resulted in a statistically significant improvement in comprehension. In general, it can be said that listeners use the time made available to them, by inserting unfilled intervals at phrase and sentence boundaries, in some way that improved their comprehension of what they heard. In subsequent experiments, the amount and distribution of pause time will be varied systematically. The information yielded by these experiments should have both practical and theoretical significance. Theoretically, it would be of considerable interest to discover those locations, within sentences, at which a listener finds processing time most useful. Such data might suggest something about the syntactic units into which the listener analyzes fluent speech. Practically speaking, a knowledge of where, in fluent speech, to insert pause time might make possible significantly greater compression of the recorded speech to which the aural reader listens.

Effects of Stimulus and Interstimulus Duration on the Immediate Recall of Time Compressed Sequences of Different Orders of Approximation to English

An experiment involving the perception of time compressed sequences of words has been performed by Mr. James Wilson. This experiment will be reported in detail in his master's thesis, and an account of it will be submitted to the Office of Education as an interim progress report. In this experiment, Wilson tested Ss for their ability to repeat sequences of words that were compressed, either by the conventional sampling method, or by removing pause time between words. The sequences of words were second, fourth, and twelfth orders of approximation to English sentences as defined by Miller and Selfridge (1950). This experiment was performed to test certain hypotheses regarding the processing operations performed by a listener as he attempts to understand spoken language.

Forward Versus Backward Reproduction of Tapes Compressed by the Electromechanical Sampling Method

Some individuals with experience in the time compression of recorded speech by Fairbanks' sampling method have reasoned that systematic differences between the onsets and the offsets of speech sounds could interact with differences in the onsets and the offsets of the samples of the original speech signal remaining after compression. One effect of such an interaction might be a more faithful reproduction of the terminal speech sounds than of initial speech sounds in syllables and words. If initial speech sounds make a greater contribution to word intelligibility than other speech sounds, it might be possible to preserve them more faithfully by reproducing the tape that is to be compressed, in the opposite direction to that used during recording. Initial speech sounds would then become terminal speech sounds.

To test this speculation, two compressed versions of a list of 100 phonetically balanced words (Egan, 1948) were prepared. These versions were identical with the single exception that the master tape used in generating them was reproduced on the compressor in the forward direction to produce one version, and the backward direction to produce the other. Each of the words was compressed to 41% of its original production time. Ten college students were divided into two comparable groups, and each group heard one of the versions. Subjects were tested one at a time, and they used earphones to listen to the test words. Each S was instructed to write, in the appropriate spaces on an answer sheet, the words he thought he heard. An intelligibility score, the number of words correctly identified, was determined for each S. These scores were distributed as follows: Forward Group -- 84, 88, 89, 84, 83; Backward Group -- 81, 81, 85, 87, 80. The difference between the means of these distributions was not significant at the 5% level. This result suggests that no advantage is to be expected by reproducing tape on a speech compressor in the backward direction.

As a final check, samples of recorded fluent speech were reproduced on the speech compressor in both directions, and the two compressed versions were compared by several judges. The superior quality of the tape reproduced in the forward direction was obvious, and the investigation was discontinued at this point.

The Experimental Control of Listening Difficulty

In those experiments on time compressed speech in which listening difficulty has been controlled or varied, Es have relied upon systematic observation for the management of this variable. That is, instead of performing operations on listening material with the intent of changing difficulty in known ways and by known amounts, they have merely examined a variety of listening materials by formulas such as the Flesch Formula (1948) and the Dale-Chall Formula (1948), and have chosen those selections that appeared to be sufficiently dissimilar or sufficiently similar for the purposes of the study.

Dr. Ronald Reid, while a graduate student at Indiana University, performed an experiment in which he attempted to gain experimental control over the difficulty variable. He measured the listening comprehension, at several accelerated word rates, of listening selections that he had varied in difficulty by rewriting them according to specified rules.

This research was reported in his doctoral dissertation. The chairman of his dissertation committee was Dr. Lawson Hughes, a member of the faculty of the Audio-Visual Center at Indiana University, who has directed several other dissertations concerned with time compressed speech. This writer was invited to serve as an ex officio member of the dissertation committee. Since the experimental question considered in the dissertation was developed through conversations involving Dr. Reid, Dr. Hughes, and this writer, and since Dr. Reid's research was similar to research proposed in Appendix A of the contract between this writer and the Office of Education for the period covered in this report, Dr. Reid was given substantial assistance in the preparation of experimental materials. He had access to project equipment, and received assistance from project staff members in the collection of data. Reid's findings, presented in his doctoral dissertation (Reid, 1968), were summarized by him, for this report, as follows.

> In order to investigate the effect on comprehension of the difficulty of material that is time compressed, an experiment was designed in which certain features of language construction that characterize "difficult" material were specifically defined and used as guides in developing "simplified" material. The comprehension tests, Forms A and B of the Nelson-Denny Reading Test, were rewritten in order to edit the language construction and make it more clear and concise. Five rules of grammar and principles of composition that characterize a high level of "readability" of material were used as guides in rewriting the material. The rewriting resulted in linguistically simplified versions of the comprehension tests. The independent variables, arranged in a 2 x 2 x 2 x 4 factorial design, were, respectively, (1) at which university the data was collected, Louisville or Indiana, (2) which of two equivalent forms of the material was used, Form A or B, (3) which of two levels of difficulty of material was used, original version or simplified version, and (4) which of four rates of presentation were used, 175, 275, 325, or 375 words per minute. The dependent variable was the number of correct responses to test questions. The inter-form reliability of the test is said to be 0.81. The analysis of covariance was used to test the statistical significance of these effects. Scholastic Aptitude Test score was the adjusting variable.
>
> The results show the following three main effects to be significant at the .01 level of significance: (A) form of material; (B) difficulty of material; (C) rate of presentation. The two following interactions were significant at the .05 level of significance: (A) university x form; (B) form x difficulty.

Main Effects

Both versions of Form B of the test resulted in greater average comprehension compared with both versions of Form A. The adjusted mean for the combined versions of Form A was 21.05, and the adjusted mean for the combined versions of Form B was 23.1.

The simplified versions of the test resulted in greater average comprehension than the original versions of the test. The adjusted mean for the original version of the test was 21.05, and the adjusted mean for the simplified versions was 23.15.

Comprehension varied significantly as a function of rate of presentation. However, the curve for the function was more or less flat until 325 and then dropped off steeply. The adjusted means for the rates of presentation were 22.65 at 175 wpm, 24.4 at 275 wpm, 22.55 at 325 wpm, and 18.75 at 375 wpm.

Interactions

The differences between the adjusted means for universities varied significantly from one form of the test to the other. The adjusted means for the Louisville group were 21.9 for Form A and 23.5 for Form B. The adjusted means for the Indiana group were 20.2 for Form A and 22.7 for Form B. Thus, the Louisville subjects on the average scored 1.6 items higher on Form B than on Form A, while this difference for Indiana subjects was 2.5.

The differences between adjusted means for difficulty levels varied from one form to the other. The adjusted means were as follows: Form A, original version 19.0, simplified version 23.0; Form B, original version 23.0, simplified version 23.1. Thus, simplifying Form A resulted in higher comprehension, while simplifying Form B had no effect on comprehension.

Pilot Studies

The Use of Filtering to Improve the Intelligibility of Speech Compressed by the Sampling Method

If a recorded speech signal is noisy, and if the noise occurs in regions of the frequency spectrum that do not contain speech information, signal

ERIC

quality may be improved simply by passing the signal through a filter that attenuates energy in the offending parts of the spectrum. In addition to this obvious application of filtering, there is some reason to believe that it may be possible to improve the intelligibility of speech signals by the use of filtering to shape the response curve in that part of the frequency spectrum containing speech information, and thus to counteract the degradation of intelligibility that results from the process of compression by the sampling method. To explore this possibility, speech signals, compressed in time by the sampling method, were passed through a Cronheit filter, which was adjusted for a variety of contours, and the resulting signals were examined aurally by several project members. If any of the filtering schemes used had resulted in an apparent improvement in intelligibility, more formal experiments would have been performed to compare the intelligibility of filtered and unfiltered speech signals. However, although different filtering schemes had discriminably different effects on voice quality, those who judged these signals detected no differences in intelligibility that would have warranted further experimentation. Consequently, this line of investigation was discontinued. It is, of course, not to be concluded that the intelligibility of speech signals cannot be improved by filtering. The experience just described suggests only that the filtering schemes tried had no apparent effect.

The Comprehension of Accelerated Speech After Prolonged Exposure

In an experiment reported in an earlier progress report (Foulke, 1964a), an evaluation of simple exposure to time compressed speech, as a means of improving its comprehensibility, was reported. This evaluation indicated that although most listeners could comprehend speech occurring at the rate of 275 wpm or less, without difficulty, and although some listeners could comprehend speech at an even faster rate, several hours of listening to accelerated speech did not improve their ability. Orr, et al., (1965) found a statistically significant improvement in comprehension test scores for Ss who listened at a word rate that was gradually increased from 325 to 475 wpm, to four full-length novels. Though these results are encouraging, the performance of Ss after prolonged exposure to accelerated speech was not as good as the performance of Ss who were tested for comprehension of listening matter presented at a normal word rate. It has already been suggested (pg. 130, ln. 23) that simple exposure may be insufficiently effective because it does not attend specifically to the acquisition of the component skills involved in the comprehension of accelerated speech. Nevertheless, the failure of simple exposure to produce the desired results may be the consequence of a failure to provide enough exposure.

During this project period, an effort was made to provide blind school children with prolonged experience in listening to accelerated speech. Students, in grades 5 through 12 at the Kentucky School for the Blind, who voluntarily engage in reading by listening for recreational purposes, were chosen as Ss. Nine children, ranging in age from 10 to 18, were found who met this requirement, and who were willing to serve as Ss in the experiment. Through consultation with the Ss, an impression of their reading tastes was formed, and books were chosen for use in the study in accordance with this impression. The experimental plan permitted each S to choose, from among the available titles, the book he wished to read by listening. His first book would be recorded with only a moderate compression in time. Upon completing his first book, each S would be invited to select a second book, and if his experience with the amount of compression represented in the first book was positive, the amount by which the recording of the second book was compressed would be increased slightly. If not, the S would be given additional experience with the initial compression. This procedure was to be followed with each S, until all Ss had read six or eight books. It was expected that, by the end of the project, Ss might be reading at rates in the neighborhood of 350 wpm. All Ss were tested for listening comprehension, before training, with one form of the STEP Listening Test in which the listening selections were presented at 350 wpm. The intention was to administer an equivalent form of the test after training and to compare pre- and post-training comprehension test scores.

There is little to report in the way of results. A few children listened to one or more books at accelerated word rates. However, the experiment was beset with mounting difficulties. Some of the tapes were so badly damaged that books had to be withdrawn from circulation among the Ss. Several Ss encountered difficulty in operating the tape recorders provided for the reproduction of tapes, and lost interest in the project. Although an effort was made to choose books that would be of general interest to the S serving in the experiment, it proved impossible to supply a broad enough selection of books to provide attractive choices for Ss with fluctuating interests. Because of these difficulties, the project was temporarily set aside.

Plans are now underway to initiate a project with similar objectives at the California School for the Blind, but with more elaborate preparation to insure its successful conclusion. Children at the school will be given substantial experience in reading by listening to accelerated speech, for both recreational and study purposes. Their ability to comprehend accelerated speech will be determined before training commences, and will be tracked during the course of training.

An Examination of the Relative Distortion of Various Speech Sounds as a Function of the Amount of Compression by the Sampling Method

When speech is compressed in time by the sampling method, brief samples of the original recording are periodically discarded, and the remaining samples are abutted in time. Since the discarding of samples is carried out on a periodic basis, there is no selectivity in respect to the portions of the speech signal that are discarded. However, if discarded samples are shorter than the speech sounds of briefest duration, some of every speech sound will be represented in the time compressed signal.

The samples discarded by the compressor used to prepare the materials discussed in this report are 40 msec. in duration. The results of Garvey (1953b), and Fairbanks and Kodman (1957), suggest that the duration of discarded samples does not affect word intelligibility substantially until it is increased beyond 40 msec. However, a discard interval of 40 msec. is long enough so that the shorter speech sounds are occasionally mutilated, depending upon the segments of a recorded tape that are not scanned by the sampling wheel on any given reproduction by the speech compressor.

Some speech sounds may contribute more to word intelligibility than others, and it would be useful to know the extent to which those speech sounds most important for word intelligibility are also the ones most likely to be mutilated by sampling accidents. Accordingly, the following investigation was undertaken.

In one study, successive compressions of single, time compressed words were compared by listeners. In one kind of comparison, a single word was compressed repeatedly, with the amount of compression held constant. It was apparent, upon listening to the successive compressed reproductions of a word, that there were variations in the signal, especially with respect to initial and final consonants. In another kind of comparison, a given word was compressed repeatedly, with the amount of compression increased for each succeeding reproduction of the word. Listening to series of words prepared in this manner suggested deteriorative changes in the quality of reproduction with increasing compression, again especially with respect to initial and final consonants.

To confirm these impressions, spectrographic records were made of the compressed words judged by listeners. The differences in successive reproductions detected by listeners, were also apparent in the spectrographic records.

As has already been pointed out, the likelihood of a sampling accident depends upon the duration of discarded samples in relation to the durations of the speech sounds that are to be sampled. At the time this investigation was undertaken, the samples discarded by the commercially available speech compressors were 40 msec. in duration, and the probability that some speech sounds would be the victims of sampling accidents was appreciable when speech was reproduced on these compressors. However, as the development of speech compression equipment continued, the duration of discarded samples was shortened, and the likelihood of sampling accidents was reduced to the point of negligibility. Since the results of this investigation would pertain only to materials produced on compressors which are now obsolete because of the improvement in compression equipment, they would have no general significance. Therefore, it was decided to terminate this line of inquiry.

Management of the Time Compression Variable

In studies concerned with the effect of compressing speech in time on its perception or comprehension, it is necessary to make a decision about the manner in which the time compression variable is to be managed. In one common approach, time compressed speech is described in terms of the average number of words spoken per minute, and word rate is varied in a linear fashion. When word rate is increased in equal steps, the fraction of original production time required for compressed reproduction decreases at a negatively accelerating rate. On a priori grounds, it has seemed to many Es that, from the viewpoint of the listener, word rate is the psychologically relevant variable, and that equal changes at the physical level might be experienced as equal changes at the psychological level.

In another common approach, the fraction of original production time required for compressed reproduction is decreased in equal steps. When this is done, the increase in word rate is positively accelerated.

Consider two hypothetical experiments. In one experiment, word rate is increased in equal steps of 35 wpm. In the other experiment, the percent of original production time required for compressed reproduction is reduced in equal steps of 10%. Table 12.1 shows the change in the percent of original production time produced by increasing word rate in equal steps, and the change in word rate produced by decreasing the fraction of original production time required for compressed reproduction in equal steps.

TABLE 12.1

ALTERNATIVE PROCEDURES FOR THE SPECIFICATION OF COMPRESSION

Words Per Minute		Percent of Original Production Time Required for Compressed Reproduction	Percent of Original Production Time Required for Compressed Reproduction		Words Per Minute
175	=	100%	100%	=	175
210	=	83%	90%	=	194
245	=	71%	80%	=	219
280	=	67%	70%	=	250
315	=	56%	60%	=	292
350	=	50%	50%	=	350

It would be useful to know those increments in word rate that produce a psychological scale of equal appearing intervals. This knowledge would provide a basis for choosing values of compression in a wide variety of experiments concerning compressed speech, and in practical applications of compressed speech. Therefore, an experiment has been planned in which, when an S places a switch in one position, he will hear a "standard word rate". When he places the switch in its other position, he will hear speech, the word rate of which he can vary by turning a control knob. During the course of the experiment, he will hear several "standard word rates" and, in each case, his task will be to adjust the variable word rate so that it matches the standard word rate. The data thus obtained should permit psychological scaling of the word rate dimension. In the case of light, we distinguish between the physical dimension of intensity, and the related psychological dimension of brightness. So, in the case of fluent speech, we may find it useful to distinguish between a physical dimension of word rate and a psychological dimension of "rapidity".

The Influence of Initial Word Rate on the Comprehension of Time Compressed Speech

When speech is compressed in time by the sampling method, there is a decline in listening comprehension. Two factors may be responsible for this decline -- a degradation in signal quality produced by the compression equipment and resulting in reduced signal legibility, and an increase in the rate at which speech sounds occur accompanied by a reduction in the duration of speech sounds. In order to gauge the relative contributions of these factors, an experiment was initiated in which three renditions of a listening selection, read at three different rates by a trained oral reader, were compressed enough to produce a final word rate of 275 wpm, and a final word rate of 325 wpm. The six resulting versions were heard by six comparable groups of Ss, who subsequently completed a multiple-choice test of listening comprehension, covering the facts and implications of the listening selection. It was hypothesized that if, at each final word rate, there was no significant difference among the three distributions of comprehension test scores, and a significant difference between the two distributions of comprehension test scores pertaining to the two final word rates, the conclusion would be that the increase in word rate was primarily responsible for the loss in comprehension. If, on the other hand, at each final word rate, there were significant differences among the three distributions of comprehension test scores, one would conclude that listening comprehension was affected by signal legibility.

Because of staffing problems, it has not yet been possible to complete the collection of data for this experiment. The three groups of Ss who heard the selection at 275 wpm were tested, and examination of their test scores reveals no significant differences in listening comprehension. On the basis of partial results, the conclusion is suggested that signal legibility, in the range in which it was varied, does not significantly affect listening comprehension. This experiment will be reactivated as soon as possible.

CHAPTER XIII

THE LOUISVILLE CONFERENCE ON TIME COMPRESSED SPEECH

by

Emerson Foulke

Abstract

The Louisville Conference on Time Compressed Speech was held at the University of Louisville on October 19, 20, and 21, 1966. The conference program included reports of experiments and demonstrations involving time compressed or expanded speech. These reports were subsequently reproduced in a volume of conference proceedings that was distributed widely. Recommendations regarding rate controlled speech were solicited from those attending the conference, and an implementation committee was appointed and instructed to act upon these recommendations. The most urgent recommendation coming from the conference was for the establishment of a center, from which it would be possible to obtain, at a moderate cost, rate controlled recorded speech of high quality, and information regarding the production, perception, and application of rate controlled recorded speech. The implementation committee acted upon this recommendation by establishing, at the University of Louisville, the Center for Rate Controlled Recordings. The implementation committee became the Board of Directors for the Center. Since its inception, the Center has responded to a steadily increasing volume of requests for information and for recordings. Under its auspices, a monthly newsletter has been prepared and is currently distributed to over 675 people each month.

The Genesis of the Conference

On December 10 and 11, 1965, Mr. Robert Bray, Chief, Division for the Blind and Physically Handicapped, Library of Congress, called together a group of people interested in exploring applications of time compressed or accelerated speech. This group recommended that a conference of national scope be held for the purpose of determining the present status of research and development with respect to the production and use of time compressed recorded speech, informing interested people of its current status, and for formulating plans relating to the future development of the area.

Accordingly, a conference was organized and presented by the University of Louisville, in collaboration with the Library of Congress. The American Printing House for the Blind, using funds made available through a grant from the Office of Education, contributed the money needed to reimburse conference participants for travel and per diem expenses. The conference was convened at the University of Louisville on October 19, 20, and 21, 1966. It was attended by approximately 100 people from all parts of the nation and from Canada, with interests ranging from the use of time compressed speech as a means of testing some aspects of cognitive theory, to the use of time compressed recorded speech in ongoing educational programs.

The Conference Program

On the first day of the conference, research reports, reports of demonstrations of the educational efficacy of time compressed speech, and demonstrations of equipment for the production of time compressed or expanded speech were presented. On the second day of the conference, conference participants were divided into seven discussion groups, and a chairman was appointed for each group. Assignment of participants to groups was made in such a way that the professions and interests represented in the conference at large were proportionally represented in each group as well. Professions represented at the conference included psychology, education, speech science, linguistics, computer science, library science, electrical engineering, school administration, and manufacturing and sales. Groups were instructed to range freely over the area in discussing the problems related to the present status of time compressed or expanded speech as a potentially useful means of communication, and its prospects for future development. They were told to have no concern for duplication of effort, in the belief that the extent of such duplication would indicate the importance of the points discussed, and that unrestricted discussion might prove more creative.

On the third day of the conference, the seven chairmen presented the assessments and recommendations of their groups. These recommendations are summarized in the section that follows. Before adjourning the conference, Mr. Bray, conference chairman, appointed an implementation committee and charged it with the responsibility of promoting the recommendations generated by the discussion groups.

Conference Recommendations

Since each discussion group was given a free hand in choosing topics to be discussed, a good deal of common ground was covered. For this reason, no effort has been made to reproduce an exact transcript of each chairman's summary. Instead, the summaries have been combined to produce a single set of recommendations.

An Economic Source for Rate Controlled Recorded Speech

The most frequent and most urgent recommendation made by conference participants was the establishment of an adequate source of supply for time compressed or expanded recorded speech. It was felt that further development of applications for rate controlled speech depends upon the organization of a center or centers capable of supplying rate controlled speech of high quality, in sufficient quantity to meet the needs of those who would use it, and at a low enough price to make its use economically feasible. It was pointed out that, as matters presently stand, it is not possible to make realistic plans for the incorporation of rate controlled recorded speech in the educational process, even for purposes of demonstration. Current costs would be prohibitive, and existing facilities could not meet the demand for the large quantities of rate controlled recordings that would be required.

Needed Research

Conference participants recognized an urgent need for further research dealing with both psychoeducational and technological problems. Many problems were mentioned that should be amenable to research. Though it will not be possible to provide a thorough statement of each problem, an effort will be made to summarize them in a general way, in the belief that such a summary may be useful to those interested in research.

The present state of ignorance regarding the nature of listening tasks, and of training methods for promoting effective listening, was felt to be a problem of central importance. It was pointed out that, because so little is known about listening of any kind, it would be a mistake to confine our research interests to just those listening tasks in which recorded speech has been accelerated. Much of what is learned about the development of listening skills may be applicable, regardless of the word rate. Presenting information at an accelerated word rate may complicate the listening task, but the impact of accelerated speech upon the perceptual and cognitive operations employed by the individual engaged in a listening task cannot be ascertained until these perceptual and cognitive operations are, themselves, more clearly understood. With such understanding, the specification of training experiences could be guided by more rational and less purely empirical considerations.

The Relationship Between Reading and Listening

A problem related to the one just discussed is the clarification of the relationship between reading and listening at both normal and accelerated word rates. Such clarification would permit more informed decisions regarding the circumstances under which accelerated listening would serve as supplementary to, or as a substitute for normal reading. Also, it would provide a basis for gauging the extent to which those procedures developed for the improvement of reading rate could be generalized to the improvement of listening rate.

Problems of Measurement

There was general recognition of the need to consider more carefully what is usually measured, and what ought to be measured in tests of listening comprehension. Researchers have, for the most part, preferred multiple-choice tests, because of their statistical reliability, ease of administration, and ease of scoring. However, such tests are valid only to the extent that they assess the factors involved in listening comprehension. It may be desirable to consider other kinds of tests, as well; for instance, tests requiring recall and reconstruction.

Another urgent problem of measurement, recognized by many, is concerned with the specification of oral reading rate. Common practice has been to specify in terms of the number of words spoken per minute. However, this approach results in considerable variability in the productions of different readers and in different productions of the same reader. One reason is that longer words require more time for their pronunciation, and are therefore produced at a slower rate. Consequently, those listening selections with longer average word rates will be read more slowly, if word rate is the measure of reading speed. Some evidence (Carroll, 1967) suggests that syllable rate provides a less variable, and more meaningful specification of reading rate. Further research on this problem is clearly indicated.

Problems of Experimental Design

Conference participants found much to criticize in the conception and design of experiments dealing with compressed or expanded speech. A frequent recommendation was that more careful attention be given to the populations sampled when Ss are recruited for experiments. It was pointed out that researchers have too often drawn their Ss from college populations, for reasons of convenience, with the hope that their results would generalize to groups such as blind school children, typical adults, and so forth. Another general criticism was that, for reasons of economy of time and effort, Es have tended to base their conclusions upon results obtained from relatively naive Ss, who were given relatively brief exposures to time compressed or expanded speech. It was recommended that experiments be performed in which the problems associated with providing prolonged exposure to time compressed or expanded speech are confronted. It was further recommended that some of these longitudinal studies involve young children, because they may be able to master very fast word rates more easily than older children or adults, just as young children can apparently master foreign languages more easily.

Organismic Variables

A host of organismic variables, the contributions of which are not well understood, were mentioned, and some were mentioned often enough and by enough people to reflect a general interest. Included were relatively unmodifiable states pertaining to basic constitution, such as mental capacity (with special reference to mental retardation) and perceptual handicaps, and relatively modifiable states such as motivation, interest, fatigue, initial resistance to accelerated speech, and attentive adjustment. The two variables mentioned last appeared to be of special interest. Many participants reported that they had sensed initial resistance to very rapid speech on the part of some listeners. They felt that an inability to overcome this resistance might limit seriously the utility of the technique and they recommended the development of procedures for overcoming this resistance. It was felt that, because of the reduced redundancy in speech compressed by the sampling method, the listener's attentive adjustment becomes a more critical problem. Normal distractions, with which the listener to normal speech has learned to contend, are likely to interfere seriously with the comprehension of accelerated speech. It was recommended that the relationship between attentive adjustment and comprehension, as word rate is increased, be given serious experimental attention.

Stimulus Variables

One frequently discussed class of stimulus variables pertained to the characteristics of the accelerated speech display. It was pointed out that aural communication may depend upon somewhat different perceptual and cognitive operations than visual communication, with the result that different vocabulary, sentence structure, format, and so forth, may be required for maximum efficiency of aural communication. It might be desirable to consider surrendering some of the time gained by the acceleration of word rate by inserting pause time at strategic points in an accelerated listening selection. Such pauses might provide needed time for implicit rehearsal, stimulus encoding, or whatever operations are involved in the process by means of which spoken language is rendered comprehensible.

It might be desirable to precede an accelerated listening selection with a list of the unfamiliar words in that selection. Presumably, this selective preview would increase the discriminibility of such words, and thus increase the likelihood of their accurate reception when they occur in the listening selection.

Since familiar selections can be understood more easily than unfamiliar selections at high compressions, it may be feasible to present listening selections for review purposes at word rates that would be much too fast for initial listening. For instance, although a word rate of 275 wpm is probably near the upper acceptable limit for initial listening, a word rate of 450 wpm might be suitable for reviewing material that has already been studied.

One of the major disadvantages of reading by listening, compressed or otherwise, in comparison with visual reading, is the reader's lack of

ERIC

control over his display. The visual reader can vary his reading rate continuously in accordance with the demands of the material being read, and can retrace with ease. He can skim through a book rapidly, and find desired information easily. The person who reads by listening, on the other hand, finds it difficult, with existing equipment, to retrace or to vary his listening rate. Finding a particular item of information in a recorded display is often quite expensive in time. It was felt that with an appropriate recorded format, involving specialized recording and playback equipment, the problems of the aural reader could be substantially reduced. For instance, if the aural reader could be provided with time compressed recorded tape, with indexing tones recorded on it at significant locations, and if this tape could be reproduced on a tape player that was variable with respect to speed and direction of tape motion, selective attention and retrieval would be greatly facilitated. If, in addition, this tape player were capable of moderate and variable compression, the disadvantages associated with aural reading could be further reduced.

Finally, it was mentioned repeatedly that the optimum speech rate would depend, in part, upon the kind of material to be heard. It was recommended that, although a beginning has been made in this regard, a good deal of research is required in order to clarify the way in which the type of listening interacts with word rate in determining listening comprehension.

Other Stimulus Variables

Stimulus variables, such as the reader's voice quality, his reading style, and natural reading rate, received frequent mention. There was also some discussion of the contribution of individual speech sounds to the intelligibility of words. It was felt that if speech sounds were affected differentially by compression in time, and if they contributed differentially to word discriminability, the interaction of these factors would have to be understood to predict the consequences of compression.

Technological Research

A strong need was felt for further development of instruments that compress or expand recorded speech by electronic or electromechanical sampling. The development of a speech compressor, with good signal quality, that can be sold cheaply enough to permit individual ownership, was regarded as an especially important objective. It was emphasized repeatedly that the current expense associated with speech compression equipment imposes a serious limitation upon the development of the area. Another insistent recommendation was for research to guide the development of playback equipment suitable for reproducing time compressed recorded speech. It was pointed out that many signal distortions, which are not critical when speech is reproduced in the original production time, may become critical with compressed reproduction. Knowledge of the effects of various kinds of distortion on the intelligibility of time compressed signals should guide the development of the equipment used to reproduce time compressed speech. The choice between earphones and loudspeaker constitutes a simple illustration. It has been found that highly compressed words are significantly more intelligible when heard over earphones instead of a loudspeaker. This is undoubtedly due to the damping problems inherent in loudspeakers that are avoided when earphones are used. Other factors to be considered in the design of a reproducer might be continuously variable control over tape speed in both directions, and the ability to record indexing signals that would be reproduced audibly at the high tape speeds used during scanning operations. Similar capability would, of course, be desirable for record reproducers. In this connection, the relative advantages of tape on open reels, tape cartridges, and records, should be examined. A study should be made to determine the feasibility of using telephone lines to distribute time compressed listening selections. For instance, a system is technically feasible in which a listener, by dialing the appropriate number, could be connected with a central facility from which he could request any listening selection in its collection and choose the word rate he preferred.

Several methods for the time compression or expansion of speech are either available or under development. Some examples are compression by periodic electromechanical sampling, compression by periodic computer sampling, harmonic compression, and compression by accelerated playback of tapes or records. These methods should be compared more carefully than they have been so far with respect to such factors as frequency response, signal distortion, word intelligibility, and listening comprehension. It is important to have this information, because the methods differ considerably with respect to such factors as cost and simplicity.

It was recommended that consideration be given to the possibility of combining methods of speech compression. The method of playing a tape or record at a faster speed than the one used during recording, though it introduces pitch distortion, has the advantage of being inexpensive and simple. Such distortion can be tolerated when compression is moderate, and this approach might be used for further tailoring the word rates of listening selections in accordance with individual preferences, that have already been moderately compressed by the more satisfactory sampling method.

Developing Uses For Rate Controlled Recorded Speech

The application of speech compression techniques to the reading problems of blind people has received considerable attention already. However, it was the general feeling of conference participants that many other uses should also be explored. It was recommended that studies be conducted to determine potential target populations for compressed or expanded speech, and that projects be organized to demonstrate the usefulness of rate controlled speech in new applications. It was suggested that there might be a considerable potential for compressed speech as a general educational tool. Compressed speech might also serve a diagnostic function in the investigation of personality or perceptual handicap. It has already shown some promise as a technique for diagnosing the underlying reason for hearing loss. Students of shorthand or typing might copy rate controlled speech that was presented initially at a very slow rate, and gradually increased in rate as their skill permitted. Expansion of the recorded speech of a user of a foreign language might be useful to a student of that language. The patient of a speech therapist might also benefit by hearing some words reproduced in more than the original production time. Mentally retarded children might, under some circumstances, receive benefit from either time expanded or time compressed speech. Many other applications may be imagined. It was the feeling of conference participants that these applications should be identified and, where feasible, developed.

Standardization of Terminology and Equipment

The lack of a standard and generally understood vocabularly of terms used in describing rate controlled speech was considered by conference participants to be a serious problem. For example, some people reserve the term "rapid speech" for describing speech that has been accelerated by reproducing a tape or record at a faster speed than the speed used during recording. To others, it has a more general significance. Similarly, to some people, the term "compressed speech", refers to speech that has been accelerated by the sampling method, while to others, it refers to speech that has been reproduced in less than the original production time, regardless of method. In describing accelerated speech, some people state the percent of compression, (either the percent of original production time saved by compressed reproduction or the percent of original production time required for compressed reproduction), and other people state the percent of acceleration. Still others state the word rate after compression, and they may or may not state the word rate before compression. It was recommended that steps be taken to arrive at a general agreement regarding the description of rate controlled recorded speech, and that some thought be given to the publication of a glossary of the terms in common use.

The need for standardization of equipment was also urged. It was pointed out that the interfacing problems arising from the lack of compatibility of recording and reproducing equipment with respect to such factors as tape speed, track configuration, response curve equalization, etc., were quite serious. Accordingly, it was recommended that an effort be made to develop equipment specifications which could serve as guidelines.

Dissemination of Information

Conference participants agreed that better publicity was needed. It was generally believed that many potential users of time compressed or expanded speech are failing to explore its possibilities simply because they are unaware of its existence. Other people, though aware, find it difficult to keep themselves informed because of the absence of a convenient source of inquiry. A variety of recommendations were made to alleviate this situation. They included the compiling of a mailing list, and the distribution of newsletters, research reports, annotated bibliographies, and demonstration tapes or records. Establishment of a speaker's bureau was also recommended. It was suggested that advantage be taken of existing dissemination facilities, such as the Educational Research Information Center (ERIC). The presentation of instructional seminars for researchers and workshops for educators and other users of time compressed speech was advocated.

Problems of Distribution

As matters presently stand, the equipment required for the satisfactory regulation of the rate of recorded speech is far too expensive for individual ownership. The only feasible alternative appears to be the establishment of a center, or centers, where economic production can be achieved. This arrangement, of course, implies some system of distribution and it was strongly urged that serious consideration be given to the orderly development of a distribution system.

The Implementation Committee

In an effort to forestall a fate that frequently befalls conferences, an implementation committee was appointed and charged with the responsibility of promoting positive action on the recommendations arising from the conference.

The Center for Rate Controlled Recordings

The implementation committee held its first meeting immediately after the close of the conference. During this meeting, plans were made for the organization of a facility that would perform two major functions: 1.) the production of rate controlled recorded speech, high in quality and low in cost; 2.) the dissemination of information about the production, perception, and use of rate controlled recorded speech. It was agreed that this facility would serve educators and researchers, primarily. As a matter of policy, it was also agreed that the facility was not to be regarded as a source for rate controlled recorded speech on a continuing basis. Rather, its function should be to stimulate the kind of experience needed to make decisions about the usefulness of rate controlled recorded speech by assisting educational institutions in organizing demonstrations involving such speech. Assistance might include the preparation of rate controlled recorded tapes and, if requested, advice concerning suitable materials, word rates, listeners, listening conditions, and experimental plans. If, by virtue of a successful demonstration, a decision was made to incorporate rate controlled speech into a school program on a continuing basis, the facility's role would be to assist the educational institution in setting up its own facilities.

After some discussion, it was agreed that this facility should be known as the Center for Rate Controlled Recordings, and that it should be located in space provided by the University of Louisville. The implementation committee then designated itself as the Center's Advisory Board, and defined for itself the role of formulating Center policy, reviewing activities engaged in by the Center, and assisting in the planning of future Center activities. This writer has served as the chairman of the Board since the Center's inception.

The Production of Rate Controlled Recorded Tapes

Since its beginning, the Center has responded to a steadily increasing volume of requests for assistance in preparing rate controlled recorded tapes for use in experiments and educational demonstrations. In some cases, recorded tape supplied by requesters has been processed at the Center to produce the desired word rates. In other cases, the Center has provided oral readers, produced recorded listening selections, and then compressed or expanded these selections in accordance with the requester's specification. To accomplish this, the Center has not only assembled the equipment required to produce rate controlled recorded tapes in any form that is likely to be requested (open reel in any conventional track configuration and playback speed, or cassette), but has also assembled the equipment needed for a recording studio of high quality.

The Dissemination of Information

Since April 21, 1967, the Center has prepared and distributed a monthly bulletin, called the CRCR Newsletter. The distribution for the newsletter has grown steadily, and it is now received by approximately 675 people.

The Center deals, by means of correspondence, with a steady stream of requests for information about rate controlled recorded speech. The Center fills a steadily growing volume of requests for research reports and demonstration tapes containing samples of time compressed and expanded speech. The director of the Center has presented discussions of rate controlled recorded speech at national conventions and conferences, and has delivered addresses concerning the production, perception, and use of rate controlled recorded speech at many schools and universities. Programs about rate controlled recorded speech have been prepared and presented on local radio and television, on national radio, and on Voice of America.

APPENDIX

1. Bowie, Walter Russell, The Story of the Old Testament. New York: Prentice-Hall, Inc., 1964. Read by: Oscar Block.

2. Chevigny, Hector, Russian America. New York: Viking Press, Inc., 1965. Read by: Kermit Murdock.

3. Clemens, Samuel Langhorn, The Adventures of Huckleberry Finn. New York: Harper and Row, Pubs., 1951. Read by: Jim Walton.

4. Cooper, James Fenimore, The Deerslayer. New York: Heritage Press, 1961. Read by: Livingston Gilbert.

5. Fromm, Erich, May Man Prevail?. New York: Doubleday & Company, Inc., 1961. Read by: Kermit Murdock.

6. Gilbreth, Frank Bunker and (Ernestine Gilbreth Carey), Cheaper by the Dozen. New York: T. Y. Crowell Co. Read by: William Gladden.

7. Hawthorne, Nathaniel, The House of the Seven Gables. New York: New American Library, Inc., 1958. Read by: Kermit Murdock.

8. Kennedy, John Fitzgerald, Profiles in Courage. New York: Simon and Schuster, Inc., 1956. Read by: Sterling North.

9. Lee, Harper, To Kill a Mockingbird. Philadelphia: J. B. Lippincott Company, 1960. Read by: Helen Shields.

10. Marshall, Catherine, A Man Called Peter. New York: McGraw-Hill Book Company, 1951. Read by: Eugenia Rawls.

11. Powell, Cyril H., Lonely Heart. Nashville: Abingdon Press, 1961. Read by: Noel Leslie.

12. Spock, Benjamin M., Dr. Spock Talks With Mothers. Boston: Houghton Mifflin Co., 1961. Read by: Paul Clark.

13. Steinbeck, John, Travels With Charley. New York: Bantam Books, Inc., 1963. Read by: Norman Rose.

14. Thoreau, Henry David, Walden (and) On the Duty of Civil Disobedience. New York: Holt, Rinehart & Winston, Inc., 1948. Read by: Kermit Murdock.

15. Wright, G. Ernest, Biblical Archeology. Philadelphia: Westminster Press, 1961. Read by: Kermit Murdock.

REFERENCES

Aaronson, D. The temporal course of perception in an immediate recall task. Journal of Experimental Psychology, 1968, 7, 129-140.

Allen, M. P. "Two Chests of Treasure", Reading Roundup Number Two. Boston: D. C. Heath & Co., 1958.

Attneave, F. Some informational aspects of visual form perception. Psychological Review, 1954, 61, 183-193.

Bellamy, M. J. An experimental study to compare the comprehension of speeded speech by blind and sighted children. (Unpub. Master's thesis, Univ. of Texas), 1966.

Bocca, E. & Calearo, C. Central hearing processes. In: J. Jerger (Ed.), Modern Developments in Audiology. New York: Academic Press, 1963.

Calearo, C. & Lazzaroni, A. Speech intelligibility in relation to the speed of the message. Laryngoscope, 1957, 67, 410-419.

Campanella, S. J. Signal analysis of speech time-compression techniques. Proceedings of the Louisville Conference on Time Compressed Speech. Louisville: University of Louisville, 1967, 108-114.

Carroll, J. B. Problems of measuring speech rate. Proceedings of the Louisville Conference on Time Compressed Speech. Louisville: University of Louisville, 1967, 88-94.

Cramer, H. L. Intelligibility of compressed speech as a function of degree and direction of delay of presentation from one ear to the other. Paper presented to APA, Chicago, September, 1965.

Cramer, H. L. The intelligibility of time-compressed speech. (Unpub. Doctoral diss., Harvard Univ.), 1968.

Cronbach, L. J. Esse·tials of Psychological Testing. (2nd ed.). New York: Harper & Brrthers, 1960.

Dale. E. & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 11-20, 28.

deQuiros, J. B Accelerated speech audiometry, an examination of test results. Translations of the Beltone Institute for Hearing Research, No. 17, 1964.

Diehl, C. F., White, R. C., & Burk, K. Rate and communication. Speech Monographs, 1959, 26, 229-232.

Durant, W. "Athenian, Spartan and Roman Life". In K. Huntress and others (Ed.), Ideas and Backgrounds. (2nd ed.). New York: American Book, 1957.

Egan, J. P. Articulation testing methods. Laryngoscope, 1948, 58, 955-991.

Enc, M. E. & Stolurow, L. M. A comparison of the effects of two recording speeds on learning and retention. The New Outlook for the Blind, 1960, 54, 39-48.

Fairbanks, G., Everitt, W. L., & Jaeger, R. P. Method for time or frequency compression-expansion of speech. Transactions of the Institute of Radio Engineers Profession Group on Audio, 1954, AU 2, 7-12.

Fairbanks, G., Guttman, N., & Miron, M. S. Auditory comprehension in relation to listening rate and selective verbal redundancy. Journal of Speech and Hearing Disorders, 1957, 22, 23-32, (a).

Fairbanks, G., Guttman, N., & Miron, M. S. Auditory comprehension of repeated high speed messages. Journal of Speech and Hearing Disorders, 1957, 22, 20-22, (b).

Fairbanks, G., Guttman, N., & Miron, M. S. Effects of time compression upon the comprehension of connected speech. Journal of Speech and Hearing Disorders, 1957, 22, 10-19, (c).

Fairbanks, G. & Kodman, F., Jr. Word intelligibility as a function of time compression. Journal of the Acoustical Society of America. 1957, 29, 636-641.

Fergen, G. K. Listening comprehension at controlled rates for children in grades IV, V, VI. (Unpub. Doctoral diss., Univ. Missouri), 1954. (Diss. Abst., 14:89, 1955).

Flesch, R. A new readability yardstick. Journal of Applied Psychology, 1948, 32, 231-233.

Fletcher, H. Speech & Hearing. New York: D. Van Nostrand, Co., Inc., 1929, 293-294.

Foulke, E. A comparison of two methods of compressing speech. Symposium Paper, Southeastern Psychological Association, Louisville, March, 1962.

Foulke, E. The comprehension of rapid speech by the blind - Part II. Cooperative Research Project #1370, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1964, (a).

Foulke, E. Transfer of a complex perceptual skill. Perceptual and Motor Skills, 1964, 18, 733-740, (b).

Foulke, E. The comprehension of rapid speech by the blind - Part III. Semi-annual Progress Report, Cooperative Research Project #2430, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1965.

Foulke, E. Comparison of comprehension of two forms of compressed speech. Exceptional Children, 1966, 33, 169-173, (a).

Foulke, E. The retention of information presented at an accelerated word rate. International Journal for the Education of the Blind, 1966, 14, (No. 1), 11-15, (b).

Foulke, E. A survey of the acceptability of rapid speech by the blind. The New Outlook for the Blind, 1966, 60, 261-265, (c).

Foulke, E. The comprehension of rapid speech by the blind - Part III. Interim Progress Report, Cooperative Research Project #2430, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1967.

Foulke, E. Listening comprehension as a function of word rate. The Journal of Communication, 1968, 18, (No. 3), 198-206.

Foulke, E., Amster, C. H., Nolan, C. Y., & Bixler, R. H. The comprehension of rapid speech by the blind. Exceptional Children, 1962, 29, 134-141.

Foulke, E. & Sticht, T. G. The intelligibility and comprehension of time compressed speech. Proceedings of the Louisville Conference on Time Compressed Speech. Louisville: University of Louisville, 1967, 21-28.

French, N. R. & Steinberg, J. C. Factors governing the intelligibility of speech sounds. The Journal of the Acoustical Society of America, 1947, 19, (No. 1), 90-119.

Friedman, H. L., Orr, D. B., Freedle, R. O., & Norris, C. M. Further research on speeded speech as an educational medium. Progress Report #2, Grant No. 7-48-7670-267, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1966.

Gabor, D. Theory of communication. The Journal of the Institution of Electrical Engineers, 1946, 93, Part III, 429-457.

Gabor, D. New possibilities in speech transmission. The Journal of the Institution of Electrical Engineers, 1947, 94, Part III, 369-387.

Garvey, W. D. The intelligibility of abbreviated speech patterns. Quarterly Journal of Speech, 1953, 39, 296-306. (a).

Garvey, W. D. The intelligibility of speeded speech. Journal of Experimental Psychology, 1953, 45, 102-108, (b).

Gerber, S. E. Dichotic and diotic presentation of speeded speech. The Journal of Communication, 1968, 18, (No. 3), 272-282.

Goldman-Eisler, F. The determinants of the rate of speech output and their mutual relations. Journal of Psychosomatic Research, 1956, 1, 137-143.

Goldstein, H. Reading and listening comprehension at various controlled rates. Teachers College, Columbia University Contributions to Education, No. 821. New York: Bureau of Publications, Teachers College, 1940.

Harris, A. J. How to Increase Reading Ability. New York: Longman Green, Co., 1947.

Hartlage, L. Differences in listening comprehension between blind and sighted subjects. International Journal for the Education of the Blind, 1963, 13, 1-6.

Harwood, K. A. Listenability and rate of presentation. Speech Monographs, 1955, 22, 57-59.

Henry, W. G., Jr. Recognition of time compressed speech as a function of word length and frequency of usage. (Unpub. Doctoral diss., Indiana Univ.), 1966.

Hutton, C. L., Jr. A psychophysical study of speech rate. (Doctoral diss., Univ. of Illinois). Ann Arbor, Mich.: University Microfilms, 1954, No. 10-494.

Iverson, L. Time compression. International Journal for the Education of the Blind, 1956, 5, 78-79.

Jester, R. & Travers, R. M. Comprehension as a function of rate and modality of presentation. Paper presented to American Psychological Association, Division 15, Chicago, September, 1965.

Johnson, W., Darley, F., & Spriestersbach, D. C. Diagnostic Methods in Speech Pathology. New York: Harper & Row, 1963, 202-203.

Klumpp, R. G. & Webster, J. C. Intelligibility of time-compressed speech. Journal of the Acoustical Society of America, 1961, 31, 265-267.

Kozhevnikov, V. A. & Chistovich, L. A. Speech: Articulation and Perception (Moscow-Leningrad). Translated by J. P. R. S., Wash., D. C., No. JPRS 30, 543, 1965.

Kurtzrock, G. H. The effects of time and frequency distortion upon word intelligibility. Speech Monographs, 1957, 24, 94.

Luterman, D. M., Welsh, O. L., & Melrose, J. Responses of aged males to time-altered speech stimuli. Journal of Speech and Hearing Research, 1966, 9, 226-230.

McLain, J. A comparison of two methods of producing rapid speech. International Journal for the Education of the Blind, 1962, 12, 40-43.

Miller, E. C. Effects on learning of variations in oral presentation. (Unpub. Doctoral diss., Univ. of Denver), 1954.

Miller, G. A. Speech and language. In: S. Stevens (Ed.), Handbook of Experimental Psychology. New York: Wiley, 1951.

Miller, G. A. What is information measurement? American Psychologist, 1953, 8, 3-11.

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

Miller, G. A. & Licklider, J. C. R. The intelligibility of interrupted speech. Journal of the Acoustical Society of America, 1950, 22, 167-173.

Miller, G. A. & Selfridge, J. A. Verbal context and the recall of meaningful material. American Journal of Psychology, 1950, 63, 176-185.

Miron, M. S. & Brown, E. R. Stimulus parameters in speech compression. The Journal of Communication, 1968, 18, (No. 3), 219-235.

Nelson, H. E. The effect of variations of rates on the recall by radio listeners of straight newscasts. Speech Monographs, 1948, 15, 173-180.

Nichols, P. G. & Stevens, L. A. Are You Listening? New York: McGraw-Hill, 1957.

Nixon, C. W., Mabson, W. E., Trimboll, F., Endicott, J. E., & Welch, B. E. Observations on man in an oxygen-helium environment at 380 mm. Hg total pressure: IV. communications. Aerospace Medicine, 1968, 39, (No. 1), 1-9.

Nixon, C. W. & Sommer, H. C. Subjective analysis of speech in helium environments. Aerospace Medicine, 1968, 39, (No. 2), 139-144.

Orr, D. B. & Friedman, H. L. Research on speeded speech as an educational medium. Progress Report, Grant No. 7-48-7670-203, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1964.

Orr, D. B. & Friedman, H. L. The effect of listening aids on the comprehension of time-compressed speech. The Journal of Communication, 1967, 17, 223-227.

Orr, D. B. & Friedman, H. L. Effect of massed practice on the comprehension of time-compressed speech. Journal of Educational Psychology, 1968, 59, 6-11.

Orr, D. B., Friedman, H. L., & Williams, J. C. C. Trainability of listening comprehension of speeded discourse. Journal of Educational Psychology, 1965, 56, (No. 3), 148-156.

Osgood, C. E. Method and Theory in Experimental Psychology. New York: Oxford University Press, Inc., 1953.

Reid, R. H. Grammatical complexity and comprehension of compressed speech. The Journal of Communication, 1968, 18, (No. 3), 236-242.

Rodgers, J. R. A formula for predicting the comprehension level of material to be presented orally. Journal of Educational Research, 1962, 56, 218-220.

Scott, R. J. Temporal effects in speech analysis and synthesis. (Unpub. Doctoral diss., Univ. of Michigan), 1965.

Scott, R. J. Computers for speech time compression. Proceedings of the Louisville Conference on Time Compressed Speech. Louisville: University of Louisville, 1967, 29-35.

Siegel, S. Nonparametric Statistics. New York: McGraw-Hill, 1956.

Sticht, T. G. Some interactions of speech rate, signal distortion, and certain linguistic factors in listening comprehension. A-V Communication Review, 1969, 17, (No. 2), 159-171.

Sticht, T. G. & Gray, B. B. The intelligibility of time compressed words as a function of age and hearing loss. Journal of Speech and Hearing Research, in press.

Taylor, E. A. Controlled Reading. Chicago: Univ. of Chicago Press, 1937, 126.

Thorndike, E. L. & Lorge, I. The Teacher's Word Book of 30,000 Words, New York: Columbia University Press, 1944.

Voor, J. B. The effect of practice upon the comprehension scores of time compressed speech. (Unpub. Master's thesis, Univ. of Louisville), 1962.

Voor, J. B. & Miller, J. M. The effect of practice on the comprehension of speeded speech. Speech Monographs, 1965, 32, 452-455.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Wood, C. D. Comprehension of compressed speech by elementary school children. Final Progress Report, Office of Education. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1965.

Woodcock, R. W. & Clark, C. R. Comprehension of a narrative passage by elementary school children as a function of listening rate, retention period, and IQ. The Journal of Communication, 1968, 18, (No. 3), 259-271.

Woodworth, R. S. & Schlosberg, H. Experimental Psychology. New York: Henry Holt & Co., Inc., 1954, 33.