TOWARDS AN IMPROVED METHOD OF PRESENTING THE LEXIPHONE CODE AND SPELLED SPEECH

by

CHING YEE SUEN

B.Sc.(Eng.), University of Hong Kong, 1966

M.Sc.(Eng.), University of Hong Kong, 1968

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in the Department of

Electrical Engineering

We accept this thesis as conforming to the

required standard

Research Supervisor ....

Members of Committee...................................

....

Acting Head of Department..............................

Members of the Department

of Electrical Engineering

THE UNIVERSITY OF BRITISH COLUMBIA

May, 1970

ABSTRACT

This thesis describes the effects of dichotic presentations on the reading speed of the users of the Lexiphone - a reading machine for the blind. The dichotic presentations investigated are: a) single delay: one signal to one ear and a delayed version of this signal to the other ear; b) multiple delay: the same signal with three interdelayed versions, two signals coming to each ear. Experiments with the Lexiphone subjects indicated that dichotic presentations (compared with ordinary binaural presentation, i.e. without delay) brought a significant improvement to their reading speed.

A similar investigation has also been made on spelled speech which has been proposed to replace the code sounds. The results indicated that multiple delay (as has been found in the case using the Lexiphone subjects) produced a little less improvement than single delay. Nevertheless, both these two dichotic presentations produced an improvement on the intelligibility of the material. The effect of the word length (number of letters contained in a word) on the intelligibility of spelled speech was also analyzed, it showed that the word length has a great effect; it was found that the percent correctness decreases with the word length. This effect also seems to be due to the longer time required to perceive the word from the spelling, thus it is suggested that a longer pause should be provided for those words with a large number of letters. Confusions of some letter sounds were observed when spelled speech was compressed, these are the consonant sounds which are articulated either at the same place or in the same manner.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF ILLUSTRATIONS	iv
LIST OF TABLES	v
ACKNOWLEDGEMENT	vi
1. INTRODUCTION	1
2. SCHEMES USED TO PRODUCE DICHOTIC SIGNALS AND TIME COMPRESSION	4
2.1 Dichotic Presentations – Methods used to Increase the Intelligibility of Sound Signals	4
2.2 Time Compression – Method used to Increase the Presentation Rate of Sound Signals	9
3. TESTS ON THE LEXIPHONE SUBJECTS	12
3.1 Testing Procedures and Results	12
3.2 Efficiency of the Dichotic Presentations	14
4. TESTS ON SPELLED SPEECH	16
4.1 Testing Procedures and Results	16
4.2 Effect of Word Length	18
4.3 Confusion of Some Letter Sounds	20
5. CONCLUSIONS AND DISCUSSIONS	22
APPENDIX I	25
APPENDIX II	27
REFERENCES	29

LIST OF ILLUSTRATIONS

Figure		Page
1	Arrangements to Produce Dichotic Signals	6
	(a) Single Delay	
	(b) Multiple Delay	
2	Computer Set-up for Processing Sound Signals	7
3	Time-delayed Waveforms	8
	(a) A Waveform and its Delayed Version	
	(b) Lower Waveform being Delayed by a Different Amount	
4	Sound Wave Compression Processes	10
	(a) Original Sound Wave	
	(b) Expansion of the Sampled Sound Wave	
	(c) New Compressed Sound Wave obtained on Playback	
5	Waveforms obtained from the Computer	11
	(a) The Original Waveform and its Expanded Version for C = 1.5	
	(b) The Original Waveform and its Expanded Version for C = 2	
6	Graph Showing the Variation of Rate Learned Versus the Rate of the Material Presented in the Lexiphone Code Test on Two Blind Subjects A and B	15
7	Combined Effect of the Rate of Presentation and the Word Length on the Percent Correctness of Spelled Speech Perceived by the Subjects	19

LIST OF TABLES

Table Page

1 Scores of Two Lexiphone Subjects on Code in Percent Correctness 13

2 Average Percent Correctness Scores of Six Subjects on Spelled Speech 17

3 Overall Percent Correctness of the Words According to the Number of Letters Contained in the Word 18

4 Letter Sounds Confused 20

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Dr. M.P. Beddoes, the supervisor of this project, for his assistance throughout the course of the work. The author also wishes to thank Dr. E.V. Bohn for reading the manuscript.

Sincere thanks are given to Mrs. M. Beddoes for preparing the spelled speech testing materials and Miss Heather DuBois for typing the manuscript.

Thanks are also due to the National Research Council for financial assistance.

1. INTRODUCTION

At present, there are two chief means of reading that provide information to the blind -- the talking book and Braille. As for printed matter, such as private documents and correspondence, newspaper, paper money, and so forth, the blind has to seek the assistance of a sighted person. Thus a machine of some sort that can read printed matter is an obvious need for the blind. In fact, for more than half a century, there has been much effort devoted to the development of such a machine. Unfortunately, up to the present, only a limited success has been attained and there is still no simple reading machine which is easy to learn and use.

Ideally, the reading machine should be easy to learn and operate, portable and cheap so that it can be carried and owned by the individual blind user. There is a gap between the first two requirements and the last. For example, the Optophone which was proposed a long time ago[1] and is still in use, is cheap and portable but difficult to learn and operate. There are recognition types of machines[2-4] which speak directly to the user, but these machines are complicated and expensive.

Although the Optophone code is difficult to learn and only a slow rate could be attained, there have been many heroic attempts by blind people to use it and it is their encouragement and the assistance of many different related organizations that provide constant zeal and assistance to the research workers in this field. The Lexiphone* is a direct translation machine which is rather more complex than the Optophone but it is believed to be easier to operate. Each letter from the Lexiphone is represented by a short sound pattern formed by both amplitude- and frequency-modulation of a square wave. This gives a melodious ouput with

*More details about the Lexiphone can be found in references 5-7.

a characteristic rhythm. Although the code melody for each letter is a tuneless one, it has been reported to be quite pleasant.

It has long been recognized that although a direct translation type of reading machine is simple and cheap compared with the recognition type, the maximum reading rate is severely limited. From the past experience, it appears to take over two hundred hours spread over a year to learn the Lexiphone code well enough to read simple sentences and stories at about thirty words per minute. Although this is a reading speed about twice that obtained with the Optophone after an equal length of training[8], it is still far behind the minimum acceptable reading rate of 50-60 words per minute set by Cooper.[9] To this problem, a method of dichotic presentation of the code sound has been proposed. This method provides an increase in the intelligibility of the code. While the mechanism by which intelligibility is improved is unclear, experimentally worthwhile increases in intelligibility have been obtained which are manifested in a useful increase in the upper limit of reading speed.

For the talking type of reading machine, practically no training time is required to understand the output sound. However, a problem arises when the machine comes across words which are not in the stored vocabulary of the computer. These words can be spelled, but this has the disadvantage of severely disrupting the listener's train of thoughts.[3]

From the length of time required to learn the code sounds of a direct translation machine and the ultimate speed attainable, and the expense and complexity of the talking machine which still has the problem of unstored words which have to be spelled out, it is worthwhile to develop an intermediate type of reading machine --- a spelled speech machine. This will reduce the amount of training time of the user and relieve him from the task of decoding the letters as it is required in the case of coding machines and yet this will not be so expensive as it is in the case of the talking machine. This can be done by using prerecorded letter sounds on a drum or disc. Each recording is suitably addressed so that it can be keyed out when required by a letter recognizer.[10] But then what reading speed can be obtained with the spelled speech type of reading machine? Are there difficulties in producing a high rate of spelled speech? Does the word-length (number of letters contained in a word) play an important role in spelled speech intelligibility? Will the dichotic presentations be applicable to spelled speech? The answers to these questions will be provided by the experiments described in later sections of this thesis.

2. SCHEMES USED TO PRODUCE DICHOTIC SIGNALS AND TIME COMPRESSION

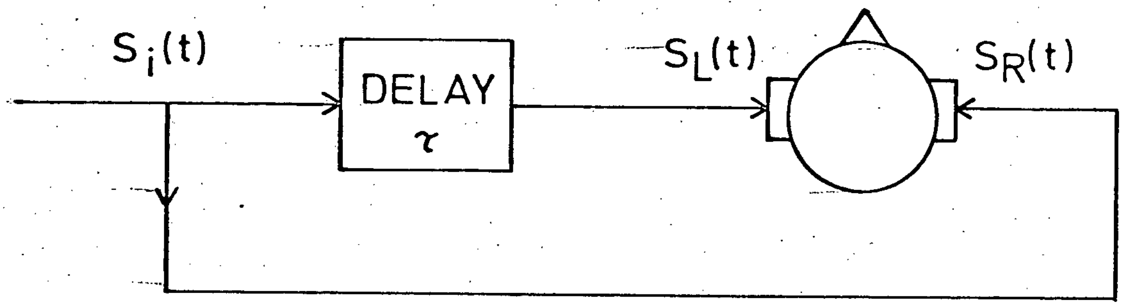
2.1 Dichotic Presentation - Methods used to Increase the Intelligibility of Sound Signals

When identical signals are presented simultaneously to both ears of a listener through a pair of earphones, a single centrally located intracranial sound image is perceived. If a time delay is introduced to one of these signals as shown in Fig. 1 (a), the single sound image is perceived to shift towards the side where the signal is leading. As the amount of delay is increased up to about 0.8 ms., then this image is perceived to be located totally either on the right or on the left according to which side the signal is leading. This kind of presentation with a time delay is known as 'dichotic' presentation and that with no delay, the ordinary binaural presentation.

Earlier research works[11-13] have shown that the binaural intelligibility of speech under a masking condition of noise can be increased by means of dichotic presentation, and the intelligibility increase brought by it has been confirmed by various groups of people[14-19] for a delay within the range of 15 ms. It has also been shown[20,21] that time delay also improves the intelligibility of rapidly compressed speech even when there is no noise. These findings seem to suggest that dichotic presentation of the above can also be applied to the Lexiphone code and spelled speech which are not so intelligible when presented at a fast rate.

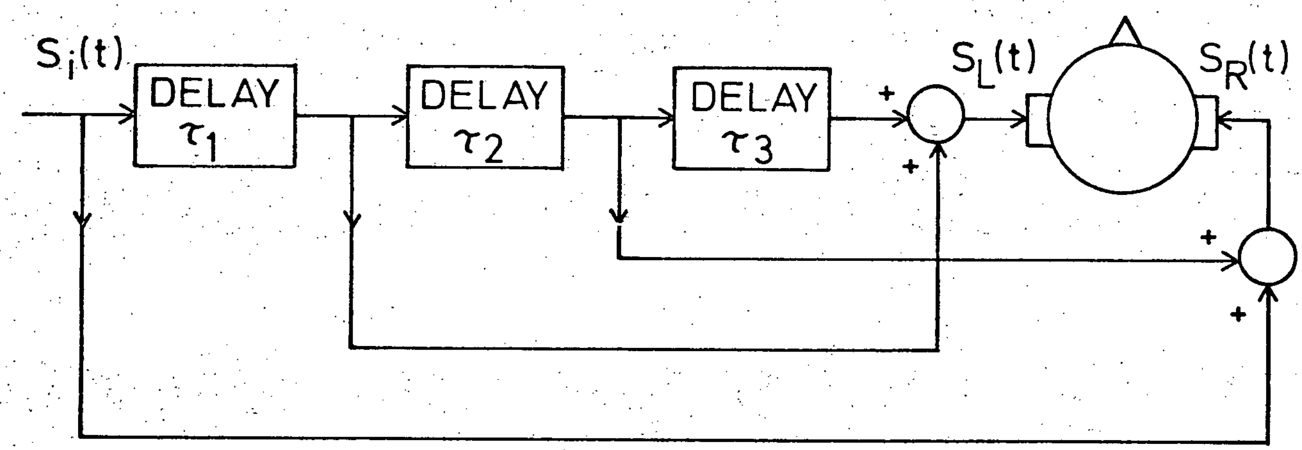
Two types of dichotic presentation have been investigated and the arrangements appear in Fig. 1 (a) and Fig. 1 (b). In Fig. 1 (a), only one signal reaches each ear. In Fig. 1 (b), two signals with a separation in time reach each ear; the purpose of the multiple delay arrangement is to test whether or not multiple time-delayed signals would produce a reinforcement effect. The required conditions of time delay on the sound signals were obtained by programming the PDP-9 computer, the set-up is shown in Fig. 2.

A program written for this purpose is shown in Appendix I. Time-delayed waveforms produced by this program are shown in Fig. 3.



(a) $S_L(t) = S_i(t-\tau)$

$S_R(t) = S_i(t)$



(b) $S_L(t) = S_i(t-\tau_1) + S_i(t-\tau_1-\tau_2-\tau_3)$

$S_R(t) = S_i(t) + S_i(t-\tau_1-\tau_2)$

Fig. 1 Arrangements to Produce Dichotic Signals: (a) Single Delay; (b) Multiple Delay.

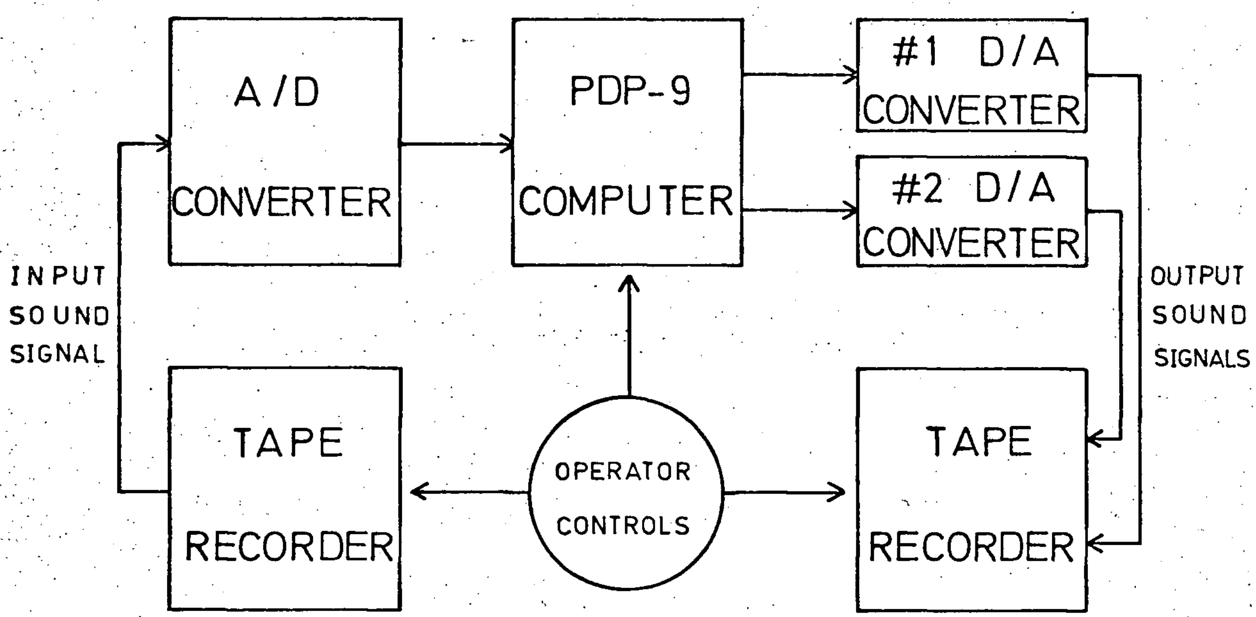
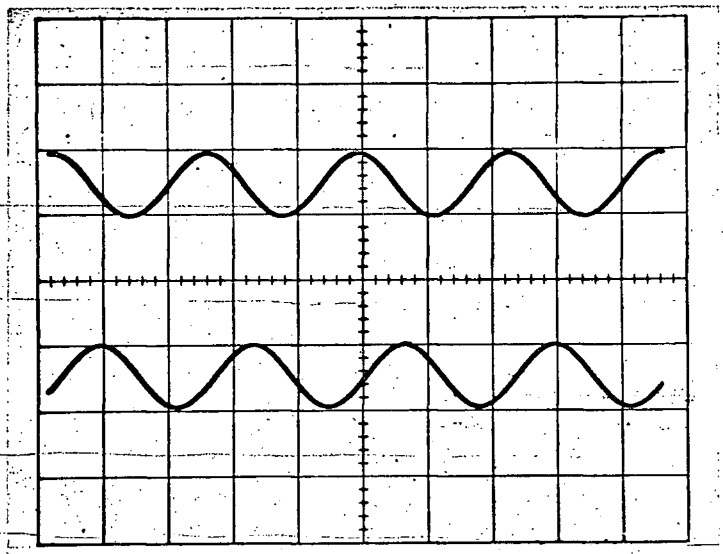
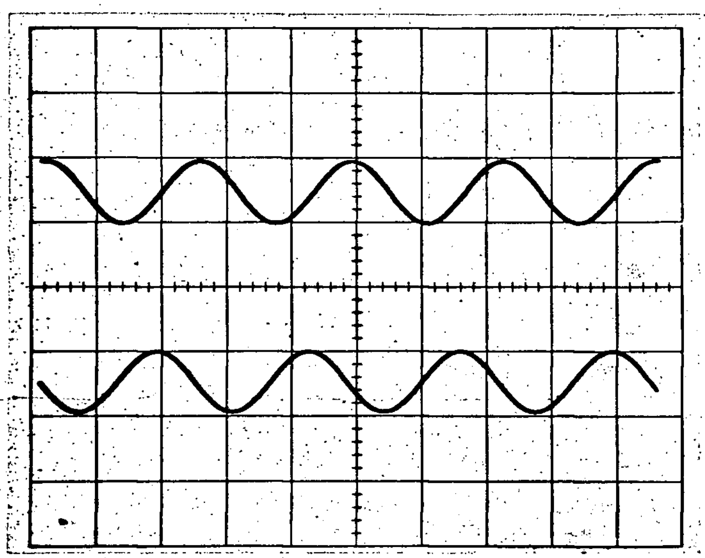


Fig. 2 Computer Set-up for Processing Sound Signals.



(a) A Waveform and its Delayed Version.



(b) Lower Waveform being Delayed by a Different Amount.

Fig. 3 Time-delayed Waveforms

2.2 Time Compression - Method used to Increase the Presentation Rate of Sound Signals

In order to obtain materials for presentation to the subjects at a fast rate, time compression on the sound signals is required. Time compression of the sound signal according to the principle of Fairbanks[22] was obtained from the PDP-9 computer using the same set-up as shown in Fig. 2. In this method, two processes are involved. The first process is to store alternate sections of the speech waveform at a sampling time interval I_s and to discard the remaining sections of the speech waveform at a discard time interval of I_d. The second process is to play the recorded waveform at an appropriate speed to restore the original pitch of the voice. These two processes are described in Fig. 4, and for the ease of illustration, equal sampling and discard intervals have been chosen (i.e. $I_s = I_d$). The compression ratio (which indicates how many times higher the rate of the compressed version of the speech sample is compared with the uncompressed speech sample) can be defined as

$$C = \frac{I_s + I_d}{I_s}$$

In the case of Fig. 4, a compression ratio of 2 has been achieved. A program is shown in Appendix II illustrating how this compression is obtained from the PDP-9 computer using a sampling interval of 30 ms.

In the spelled speech experiment, two compression ratios (1.5 and 2) were used to bring the original material to 1.5 and 2 times faster. Both the original and the resulting expanded waveforms coming out from the D/A converters are shown in Fig. 5 (a) and Fig. 5 (b). The disconnected lines indicate the place where part of the waveform has been discarded. Time compression is obtained by playing the recorded waveforms respectively at 1.5 and 2 times faster.

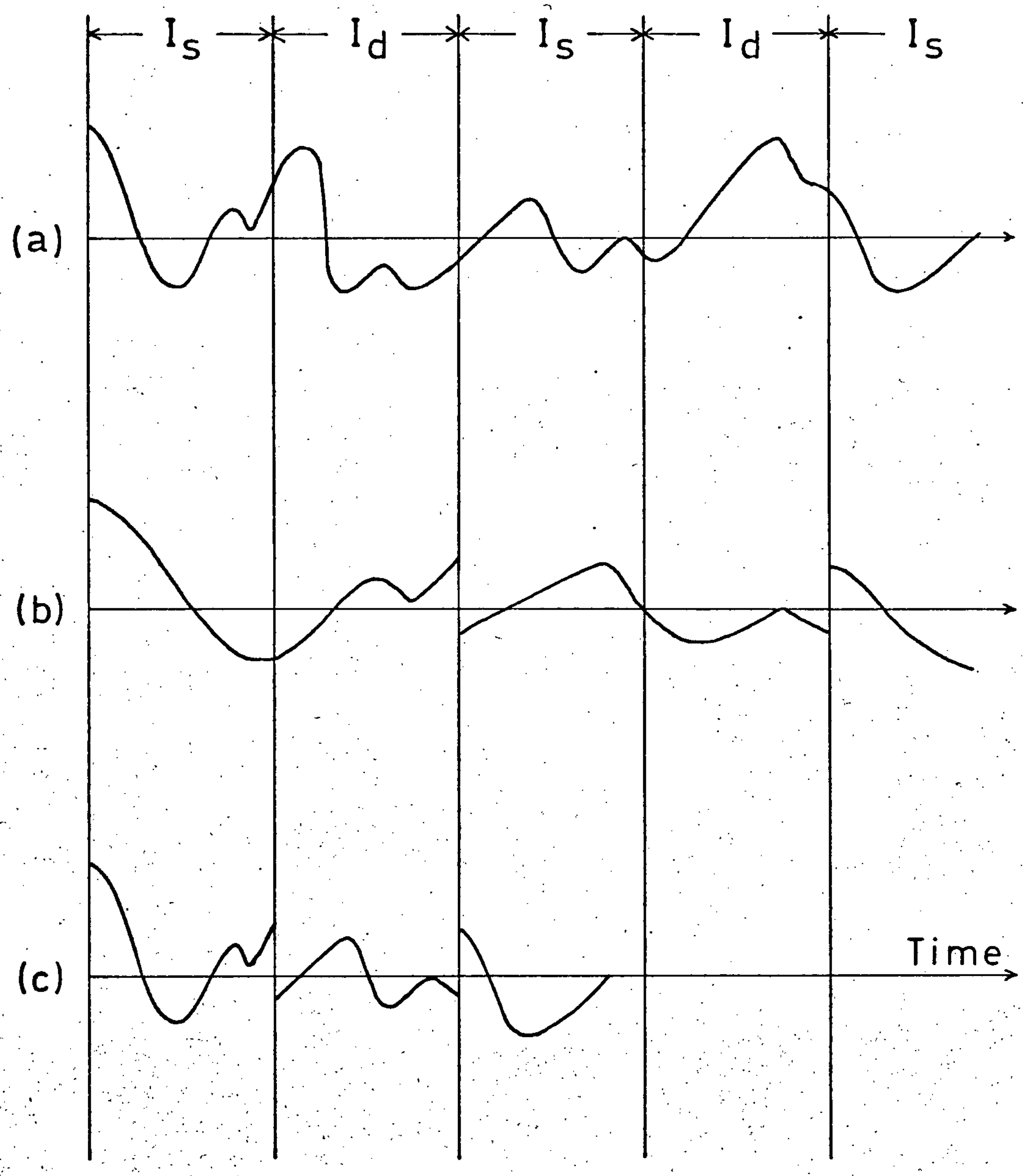
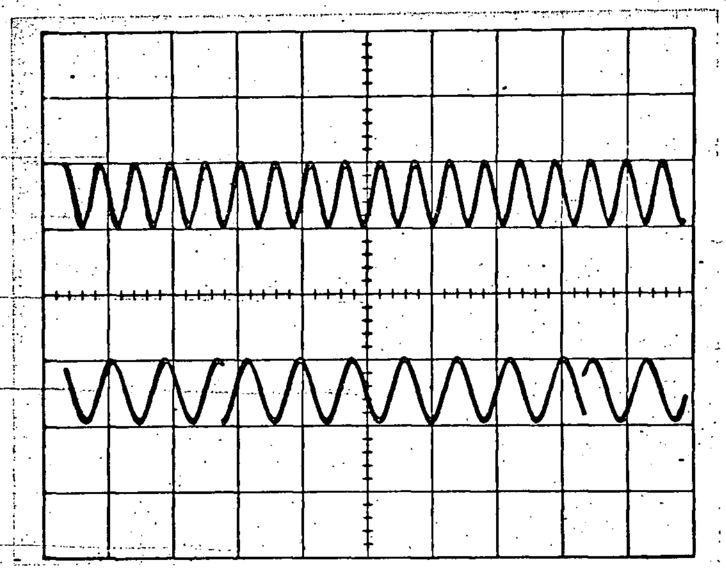
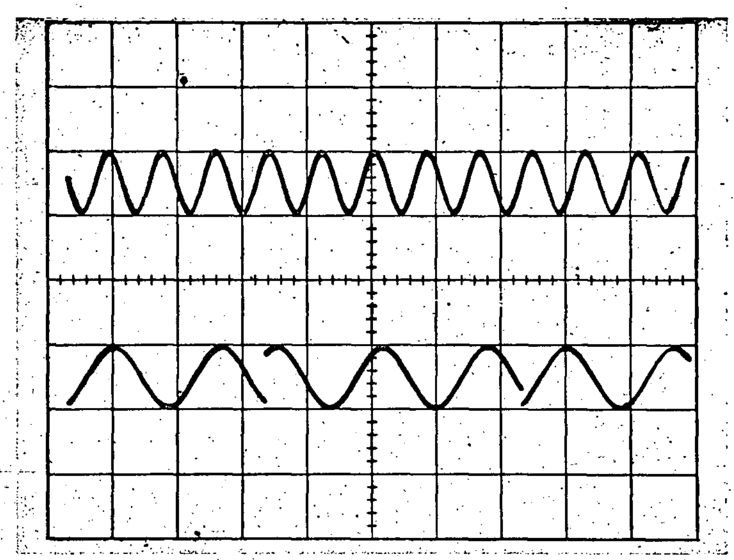


Fig. 4 Sound Wave Compression Processes: (a) Original Sound Wave; (b) Expansion of the Sampled Sound Wave; (c) New compressed Sound Wave obtained on Playback.



(a) The Original Waveform and its Expanded Version for $C = 1.5$.



(b) The Original Waveform and its Expanded Version for $C = 2$.

Fig. 5 Waveforms obtained from the Computer.

3. TESTS ON THE LEXIPHONE SUBJECTS

3.1 Testing Procedures and Results

In order to see how effective the dichotic presentations are compared with the ordinary binaural presentation of the Lexiphone code under various reading speeds, tests were performed on two Lexiphone subjects. To ensure that the test materials are reasonably uniform in difficulty, a long story[23] with a limited vocabulary of the first thousand words of the Thorndike-Lorge lists was chosen. These materials were first processed on a computer to give the desired S_L and S_R signals (as shown in Fig. 1 (a) and Fig. 1 (b)) at the tested speeds and were then recorded on a tape recorder. Tests were conducted in the following presentations:

1) Ordinary binaural presentation, i.e. when there is no delay or $\tau = 0$ ms.

2) Dichotic presentations:

a) Single delay: $\tau = 0.8$ ms. and

b) Multiple delay: $\tau_1 = 0.25$ ms., $\tau_2 = 0.3$ ms. and $\tau_3 = 0.25$ ms.

A time delay of 0.8 ms. (which represents the longest delay experienced in daily life when the sound source is either on the right hand side or on the left hand side of the ears) had been chosen because it has not yet been shown that a greater amount of delay will give additional benefit. Besides, this amount of time delay can be implemented easily employing an audio delay line.

The passages were recorded at three different speeds corresponding to 30, 40 and 50 words per minute. The rates of 30 and 40 words per minute represent the peak reading rates of the Lexiphone subjects. The processed materials were presented in random order. The tests were carried out in a quiet room and the subjects listened to the recorded materials at a comfortable level through a pair of headphones; they were instructed to omit the words if they could not catch them. After an explanation of the testing procedures and a run of four practice passages, the actual tests were carried out. The scores were based on the response of 90 words in the middle of the passages each of which consisted of about 120 words. The scores of these tests are shown in the following table with each figure representing the mean of five tests. In view of the good results of subject A, an additional test was performed on her at a rate of 60 words per minute. The score was: 72.2% for no delay and 98.9% for both dichotic presentations.

Conditions / Speed(w.p.m.)	1 No Delay		2 Single Delay		3 Multiple Delay	
	Subject A	Subject B	Subject A	Subject B	Subject A	Subject B
30	82.9	57.8	96.7	64.3	97.1	62.9
40	90.2	61.7	98.9	73.1	96.5	72.2
50	83.1	58.2	96.2	66.2	98.9	62.9

Table 1 Scores of two Lexiphone Subjects on Code in Percent Correctness

Note: At the time of this test, subject B had been away from the code for several months.

Statistical analysis of the data by means of the analysis of variance revealed that the dichotic presentations have definite levels of significance. The analysis showed that there are highly significant differences ($p < 0.1\%$) among the different presentations and there are also significant differences ($p < 5\%$) among the speeds of presentation.

By comparison, the results in column 2 and column 3 of table 1 show significant improvement. With dichotic presentation, the subjects' scores are higher than using normal binaural presentation at all reading rates. It also indicated that multiple delay is only about as effective as a single delay; this may be due to some confusion caused by a repetition of the same signals even though they are separated apart in time.

3.2 Efficiency of the Dichotic Presentations

Fig. 6 shows the variation of the rate of Lexiphone code information learned versus the rate presented in the three tested conditions. The graphs of the dichotic presentations are seen to be closer to the ideal case of perfect learning, the effect is more pronounced in the performance of the better subject (subject A).

Thus message efficiency, as measured by the number of words learned per unit time, increases with the dichotic presentations. Message efficiency starts to decline at the rate of about 40 words per minute for the binaural presentation.

Judging from Fig. 6, subject A can read with the Lexiphone more efficiently at a high rate; for subject B, because of the decline around the region of 40 w.p.m., this is probably the most efficient rate of presentation for her.

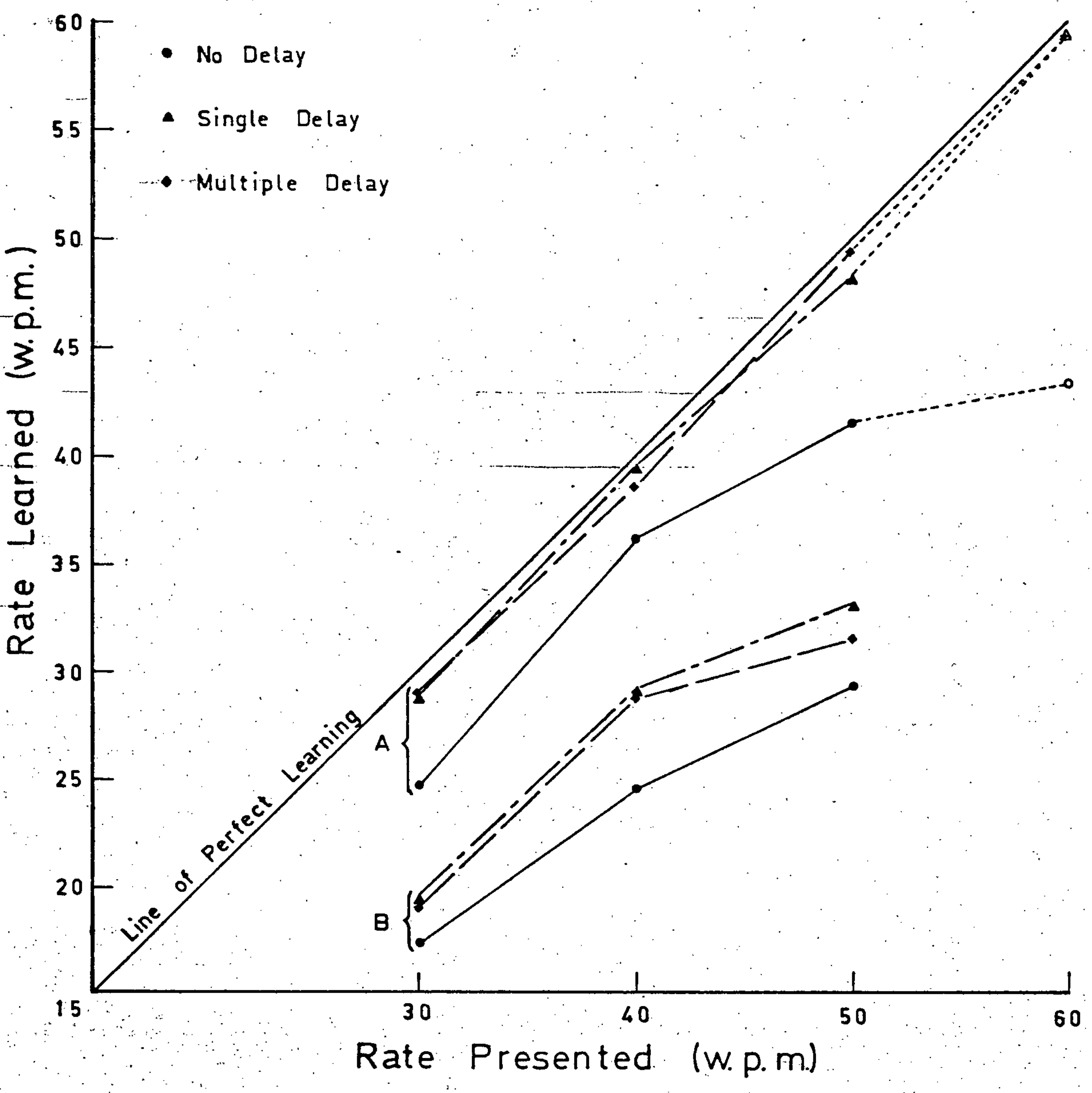


Fig. 6 Graph Showing the Variation of Rate Learned Versus the Rate of the Material Presented in the Lexiphone Code Test on Two Blind Subjects A and B.

4. TESTS ON SPELLED SPEECH

4.1 Testing Procedures and Results

In order to find out the efficiency of spelled speech as a medium of communication and test the effect of dichotic listening compared with the ordinary binaural presentation, tests were performed. The test material consists of selected lists of phonetically-balanced sentences.[24] Each list has an average length of seventy-eight words and consists of ten sentences. These lists were spelled by a mature female speaker at an average rate of fifty-four words per minute. The material was then compressed on the PDP-9 computer to give three different rates for presentation, i.e. 54, 81, and 108 words per minute respectively. These materials were again processed on the PDP-9 computer to obtain the three different kinds of presentation, i.e. no delay, single delay and multiple delay as it was done in the Lexiphone code test. Six subjects (three male and three female university undergraduate students) who are naive with respect to spelled speech and compressed speech were employed in this experiment. Both the material and the conditions of presentation were randomized in a balanced design so that all the testing lists were covered by all the subjects, all the rates of presentation and all the listening conditions (including the three presentations: no delay, single delay and multiple delay and also the interchange of the signals to the right ear and the left ear). The tests were carried out in a quiet room and the subjects listened to the recorded materials at a comfortable level through a pair of headphones; they were instructed to omit the words if they could not catch them. Before the actual test began, the testing procedure was explained to the subjects and each subject had a practice of five lists. The scores of these tests are shown in Table 2.

Speed (w.p.m.) \ Conditions	1 No Delay	2 Single Delay	3 Multiple Delay
54	77.66	77.68	78.33
81	63.81	72.58	68.65
108	56.69	62.41	63.98

Table 2 Average % Correctness Scores of Six Subjects on Spelled Speech

On analyzing the result of the spelled speech experiment, it has been found that the effects of the different presentations are significant ($p < 5\%$), the rate of presentation of the material is highly significant ($p < 0.1\%$). The analysis also revealed that interchanging the signals of the two ears has no significant affect at all ($p > 50\%$), this indicates that it does not matter whether the right ear signal leads the left ear or vice versa.

On comparing the results shown in Table 2, it can be seen that dichotic presentations give improvements at the three tested speeds of presentation. However, at the lowest rate of presentation (54 w.p.m.), there is only a minor improvement. This may be due to the fact that at this low rate, it is not the intelligibility that affects the performance, but it is the difficulty of putting the letters into words. It is expected that a trained subject on spelled speech should have no difficulty at all in obtaining a high score ($>90\%$) at this rate of presentation. The results at the medium and the fastest rates seem to resemble the results of subject B in the Lexiphone code test. If it is the subject's skill (such as subject A in the code test) that makes dichotic presentations more helpful, then it is very likely that subjects trained in spelled speech will be more beneficial from the dichotic presentations. This has to be confirmed after training a group of subjects on hearing spelled speech.

4.2 Effect of Word Length

During the test, a large number of errors occurred in those words which contained a large number of letters. In view of this, a calculation was made based on the number of errors observed at different word length (the number of letters contained in a word). The overall picture of the effect of word length of the entire experiment is shown below.

Word Length (letters)	3 or less	4	5	6	7	8
% Correctness	86.22	66.74	51.84	43.13	29.67	32.04

Table 3 Overall Percent Correctness of the Words According to the Number of Letters Contained in the Word.

The combined effect of the rate of presentation and the word length appears in Fig. 7. From Table 3 and Fig. 7, it can be seen that the word length has a very great effect on the correctness of the response. The longer the word, the more likely that an error will be made. It must be pointed out that no longer pause was allowed for long words in the test. This may create the situation of insufficient time for the perception of the long words from the letters, and while the subject was still pondering on the long words, words of the subsequent order followed and disruption occurred. Thus it is suggested that for those words with four or more letters, a slightly longer pause should follow to compensate for the time required to perceive these words from their spelling.

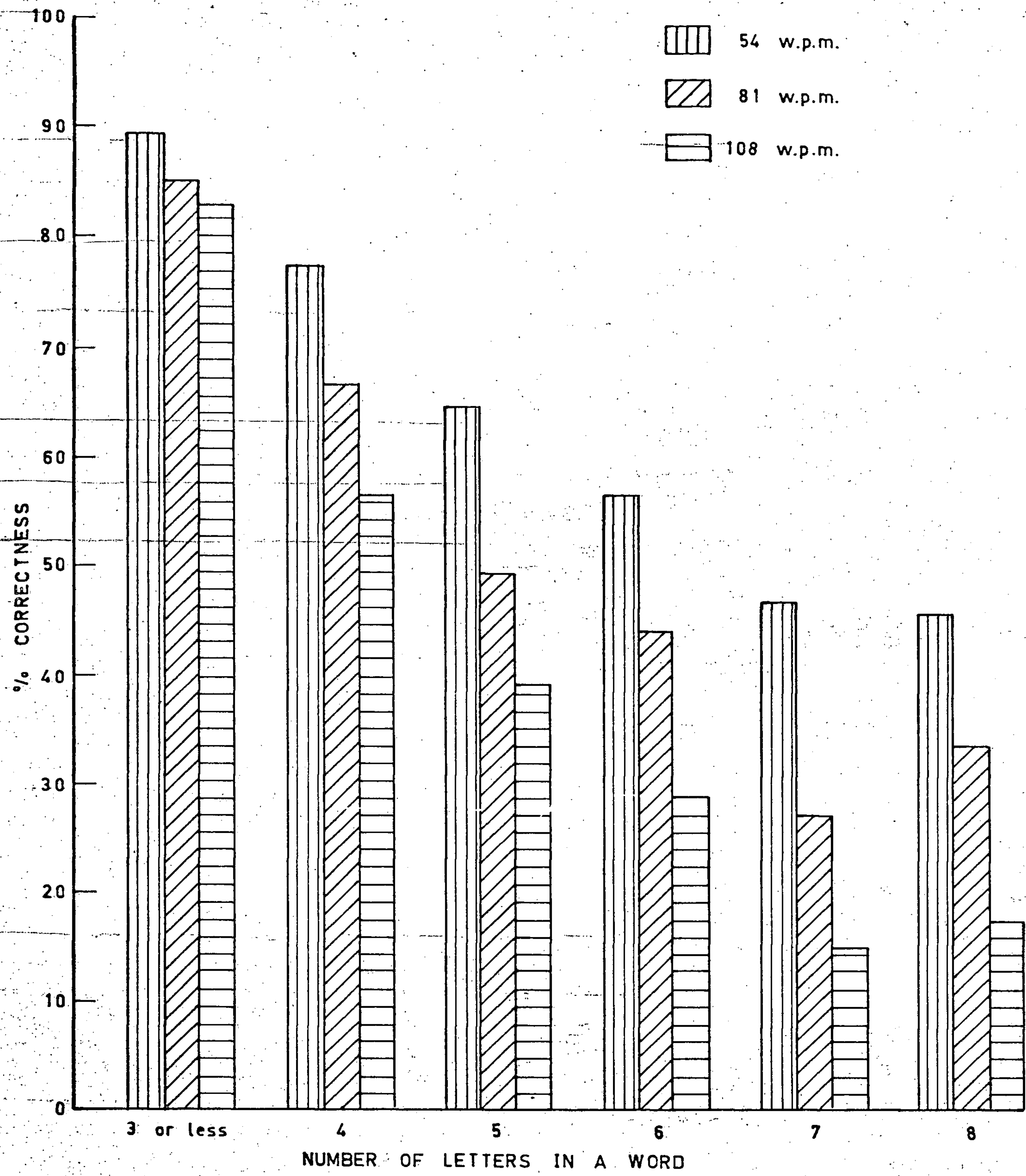


Fig. 7 Combined Effect of the Rate of Presentation and the Word Length on the Percent Correctness of Spelled Speech Perceived by the Subjects.

4.3 Confusion of Some Letter Sounds

During the spelled speech test, confusion of some letter sounds was observed in the case of compressed materials. The total number of errors due to the confusion of letter sounds is about 5.5%. All the confused letters are consonants. Confusion at the lowest speed (uncompressed) was unnoticeable. The following table shows how the letter sounds were confused according to the manner and the place of articulation of the letter sounds.[25]

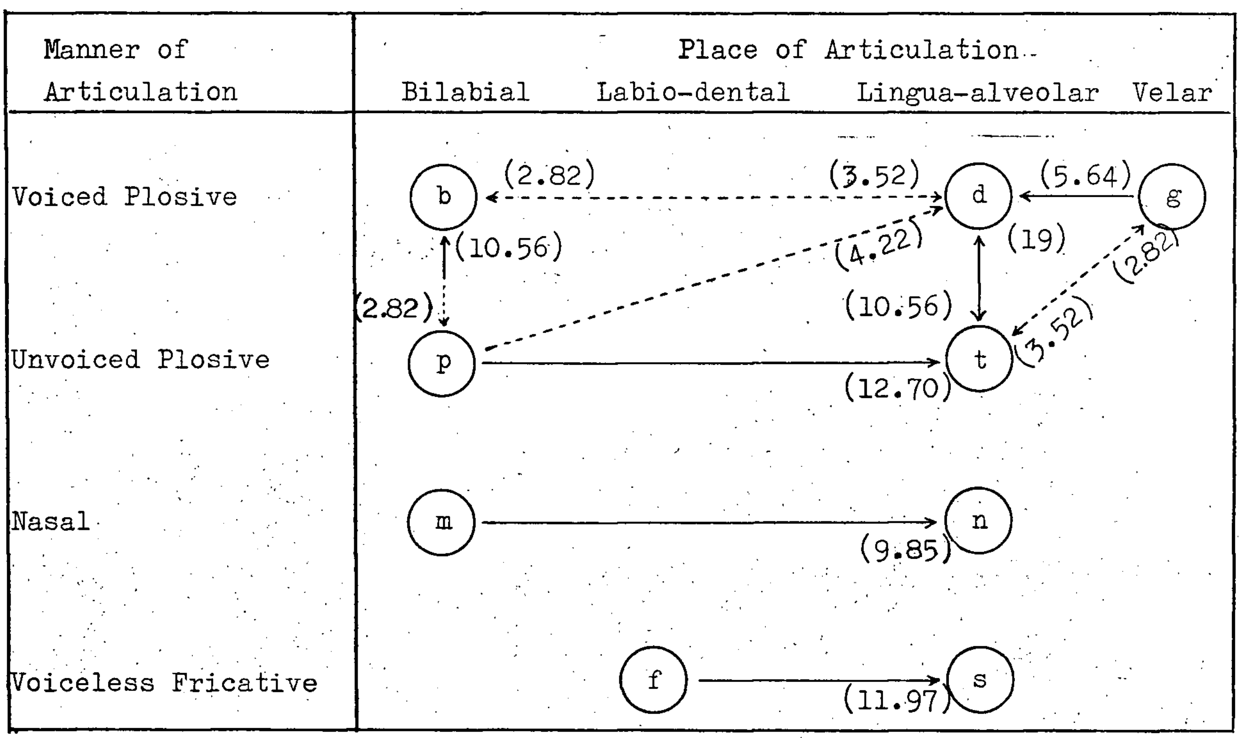


Table 4 Letter Sounds Confused. (Arrow head indicates the direction of confusion, e.g. the actual sound 'm' is mistaken to be the sound 'n', and the sound 'd' is mistaken as the sound 't' and vice versa. A dotted line indicates a less frequent confusing situation. The figures in brackets denote the percent contribution of each confusing situation towards the total number of confusions.)

Table 4 indicates that the plosive sounds are the mainly confused sounds. It also indicates that letter sounds articulated in the same manner or at the same place are easier to be confused. There are some uni-directional confusions which indicate that some of them can be mistaken as others, but not vice versa.

The reason that this kind of confusion occurred only in consonants with the compressed materials and not with its uncompressed version is because the duration of phonation (the length of time involved in the production of a single sound) of the consonant sounds is much shorter than that of the vowels[26,27], the ratio is roughly 3 to 2 (the duration of phonation for vowels averages to 0.117 sec. and the duration of phonation of consonants averages to 0.08 sec.[27]). During the compression process, a part of the letter sound has been discarded effecting a shorter duration of the letter sound. This makes the compressed consonant sounds more difficult to distinguish (especially those which are articulated in the same manner or at the same place) and thus confusion occurred.

Although confusions of some letters may be eliminated by correct guessing from the context of the material, easy identification of letter sounds in spelled speech is important because some words consist of single letters and differentiation of word patterns would be facilitated if there were no confusions among letters. Thus it appears desirable to substitute special sounds for the more frequently confused letters such as d, t, p, m and f.

5. CONCLUSIONS AND DISCUSSIONS

From the Lexiphone and spelled speech experiments, it has been found that both the single delay and multiple delay dichotic presentations bring significant improvements to the intelligibility of the materials presented through the earphones. The Lexiphone code tests indicate that dichotic presentations are more helpful to the subject who is better skilled in the code. The results of the spelled speech tests resemble that of subject B (the subject who is not so skilful in using the Lexiphone) in the Lexiphone test. If it is the skill of the subject which had produced this strong influence, then dichotic presentations would be more beneficial to those who have been trained on spelled speech. This has to be confirmed on experimenting subjects who have already been trained on spelled speech. Conclusions on this factor may also be arrived using other better skilled Lexiphone users. Since it takes a rather long time (over 200 hours of training) to learn the code well, it would be more desirable to train subjects on spelled speech for this purpose. In fact, it has been found[28], that after a training time of only about twenty hours, if words were spelled rapidly enough, the whole word, rather than the individual letters, would become the unit of perception.

From the comments made by the spelled speech subjects, it indicates that a better performance may be arrived at with more practice. The subjects also remarked that the fastest rate of presentation (108 w.p.m.) was too fast for the perception of the words, but they all felt that this situation might change after they had been exposed to spelled speech for a long enough time. The subjects also showed no particular preference to any of the listening conditions.

Counting the percent correctness of the subjects' performance in words formed from different number of letters indicates that the longer the word, the easier errors were made, and while they were still pondering on the long words, other words that followed disrupted the decoding process of the long words. Thus it would be helpful if a longer pause is provided for those words containing a bigger number of letters.

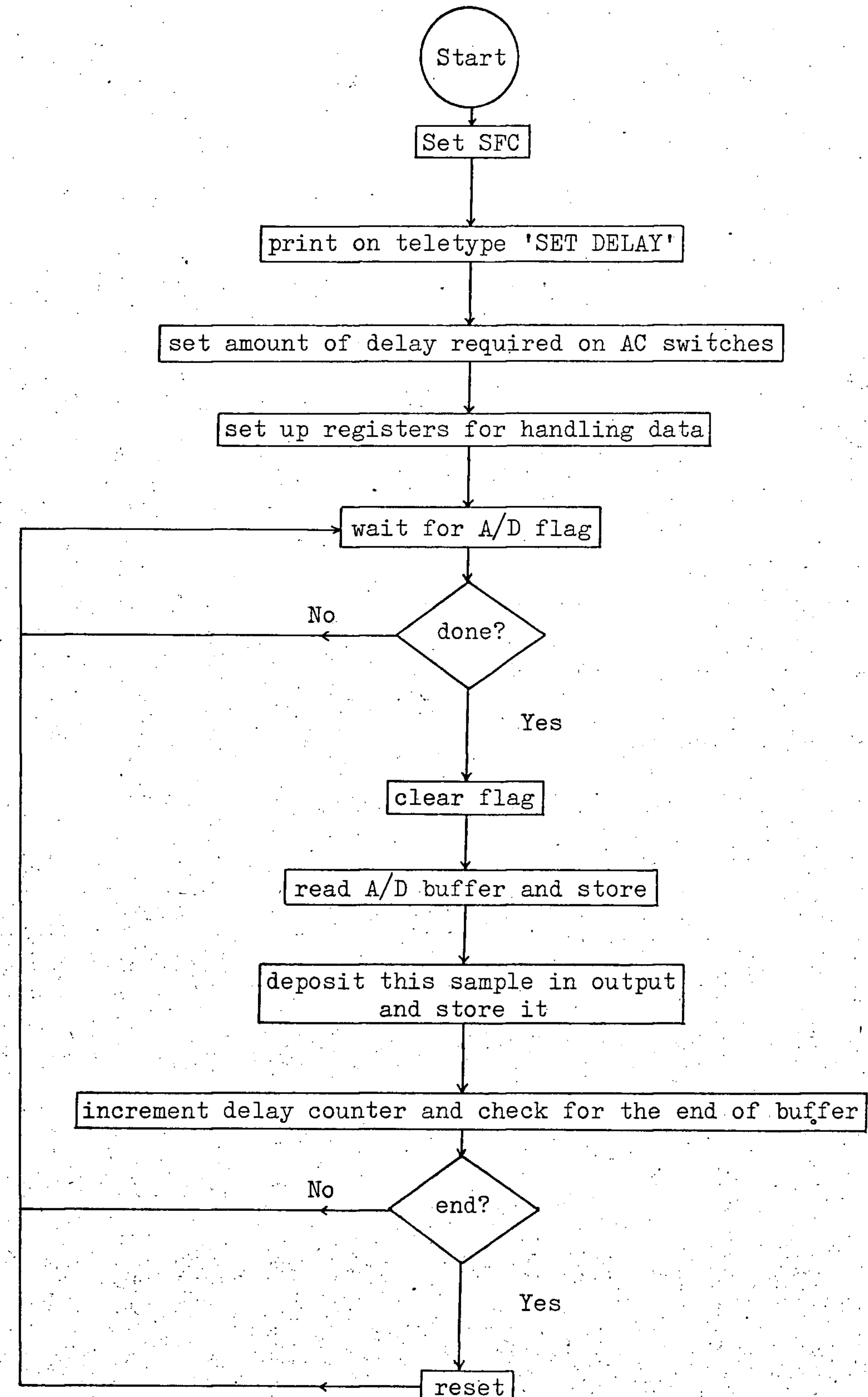
Another finding of the spelled speech test is that there are errors caused by the confusion among the letter sounds, in particular, the consonants. In order to bring spelled speech into a more useful medium for reading machine work, the observed confusions should be removed. This can be done by substituting these letter sounds with special sounds, but this has the disadvantage of learning the special sounds. An alternative is to lengthen the duration of phonation of these sounds so that they become more distinguishable. A similar effect can be achieved by introducing some other variables, such as loudness, accent and intonation. It would be very useful to find out the minimum duration of phonation for the perception of the different letter sounds. One way of doing this is to simulate the letter sounds on the computer and then control the duration of the letter sounds for presentation. This will not only find out the duration of phonation required for the perception of the letter sounds, but also give an idea of the absolute maximum rate of presentation of spelled speech.

In this thesis, only two dichotic presentations have been investigated and the effect of the sampling period of compression has not been considered at great depth. It must be pointed out that there are other dichotic presentations which are worthwhile for further research: these include a longer single delay (in the range of 20 ms. to 300 ms.) and dichotic compression (presenting the samples in the sampling interval to one ear and the samples in the discarded interval to the other ear so that there is no loss of information to the ears). The sampling interval is an important factor affecting the intelligibility of compressed speech because it is related to the amount of information retained from the originally uncompressed information for presentation to the ears. Although this has been investigated by several research workers in this field[20-22,29] there is still the lack of a model for theorizing the optimum sampling interval for compression. An investigation of this sort would be very useful to the field of compressed speech. It is likely that the optimum sampling interval is related to the duration of phonation of different letter sounds, the compression ratio and the pitch of the voice. When this model has been established, it will be easier to produce intelligible letter sounds at different speeds of presentation for the spelled speech reading machine and it will also throw some light onto better methods of dichotic presentations.

Appendix I

FLOW CHART FOR THE DELAY PROGRAM

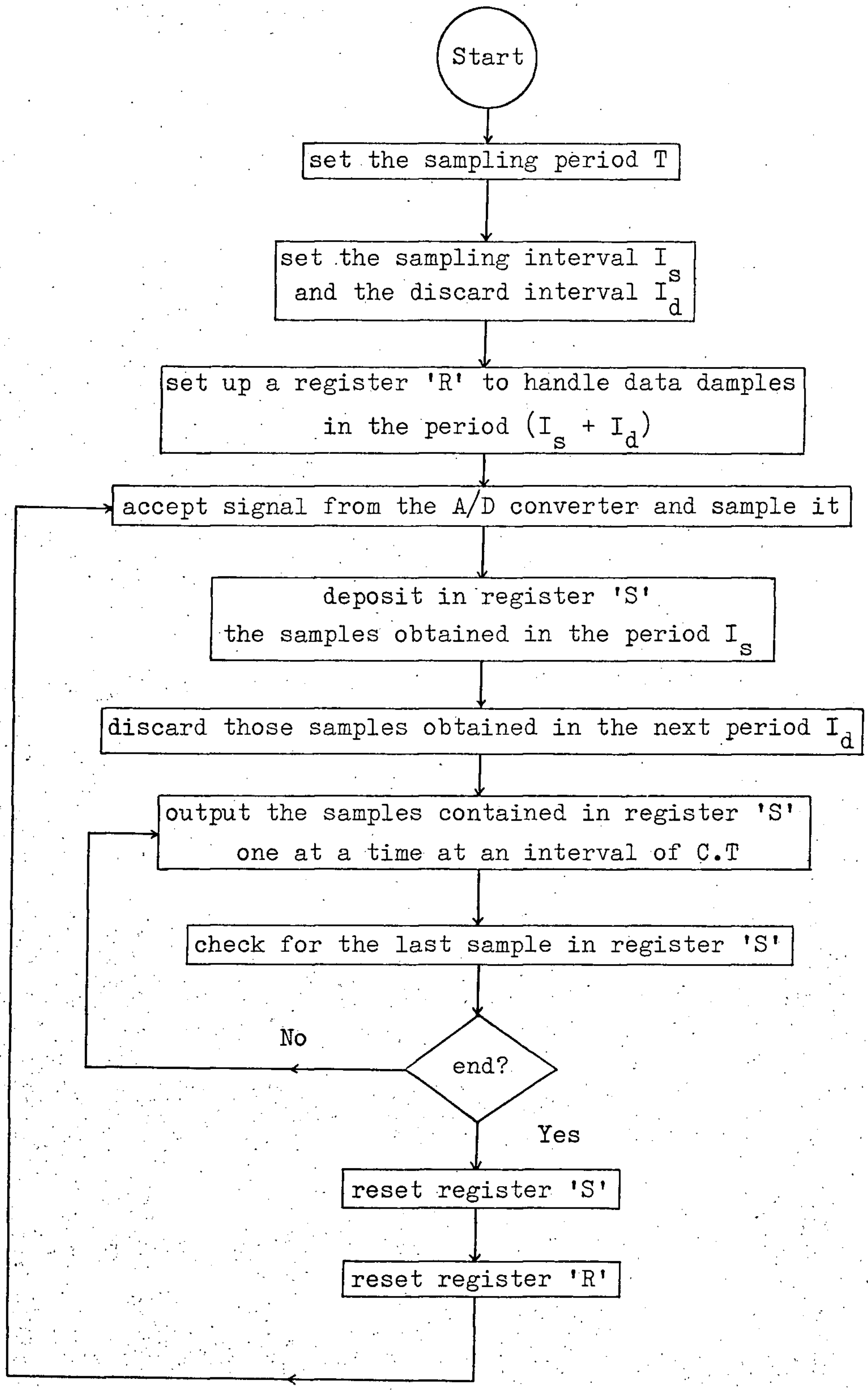
The following is a simplified flow chart for the delay program.



```
        .TITLE DELAY
        .IODEV 10
SFC=701207
IN52=705217
OUT31=703107
OUT32=703207
SKPFLG=705301
CLRFLG=705302
START   .INIT 10,1,START
        .WRITE 10,2,PRINT,3
        .CLOSE 10
        LAC (-62
        SFC                     /SET THE SAMPLING PERIOD = 50 US
        HLT                     /HALT, SET THE AMOUNT OF DELAY
        LAS                     /ON AC SWITCHES, PRESS CONTINUE
        CMA
        TAD (1
        DAC DEL#
        LAC (DATA-1
        DAC* (10
        DAC* (11
ONE     SKPFLG
        JMP .-1
        CLRFLG
        IN52
        OUT31
        DAC* 10
        ISZ DEL
        JMP ONE
TWO     SKPFLG
        JMP .-1
        CLRFLG
        IN52
        OUT31
        DAC* 10
        LAC* 11
        OUT32
        LAC* (10
        SAD (DATA+16660
        JMP RESET1
        LAC* (11
        SAD (DATA+16660
        JMP RESET2
        JMP TWO
RESET1  LAC (DATA-1
        DAC* (10
        JMP TWO
RESET2  LAC (DATA-1
        DAC* (11
        JMP TWO
PRINT   003002
        0
        .ASCII 'SET DELAY'<15>
DATA    .BLOCK 16665
        .END START
```

Appendix II

FLOW CHART FOR TIME COMPRESSION PROGRAM



```
        .TITLE  TIME COMPRESSION, C=2
        .IODEV 10
SFC=701207
IN52=705217
OUT31=703107
OUT32=703207
SKPFLG=705301
CLRFLG=705302
/
START   .INIT 10,1,START              /PRINT OUT TITLE
        .WRITE 10,2,PRINT,10
        .CLOSE 10
        LAC (-62
        SFC                           /SET THE SAMPLING PERIOD = 50 US
RUN     LAC (DATA-1
        DAC* (10
        DAC* (11
NOW     SKPFLG
        JMP .-1
        CLRFLG
        IN52
        OUT31
        DAC* 10
THERE   SKPFLG
        JMP .-1
        CLRFLG
        IN52
        OUT31
        DAC* 10
THEN    LAC* 11
        OUT32
        LAC* (10
        SAD (DATA+2257
        JMP RUN
        JMP NOW
PRINT   010002
        0
        .ASCII 'TIME COMPRESSION, C=2'<15>
DATA    .BLOCK 2270
        .END START
```

References

1. d'Albe, F. "The type-reading Optophone", Nature, Vol. 94, p. 4, 1914.

2. Cooper, F.S. and Gaitenby, J. "Reading Machine Research at Haskins" in 'Reading Machines for the Blind, Sixth Technical Session' held at the Veterans Administration Central Office, Washington, D.C. in Jan. 1966.

3. Cooper, F.S., Gaitenby, J., Mattingly, I.G. and Umeda, N. "Reading Aids for the Blind: A special case of Machine-to-Man Communication", IEEE Trans. Audio and Electroacoustics, Vol. AU-17, No. 4, 266-269, Dec. 1969.

4. Lee, F.F. "Reading Machine: from Text to Speech", IEEE Trans. Audio and Electroacoustics, Vol. AU-17, No. 4, 275-282, Dec. 1969.

5. Caple, G. "The Lexiphone, a simple Reading Machine for the Blind," M.A.Sc. Thesis, U.B.C., January 1966.

6. Beddoes, M.P. "An Inexpensive Reading Instrument with a Sound Output for the Blind", IEEE Trans. Bio-Med. Eng., Vol. BME-15, 70-79, April 1968.

7. Ramsey, D. "Design of a Simple Reading Machine for the Blind", M.A.Sc. Thesis, U.B.C., October 1968.

8. Coffey, J.L. and McFarland, R.R. "The Evaluation and Standardization of Selection and Training Procedures for the Battelle Aural Reading Device" Research Report to the Veterans Administration from the Battelle Memorial Institute, Columbus, Ohio, June 1963.

9. Cooper, F.S. "Research on Reading Machines for the Blind" in P.A. Zahl (ed.), Blindness: Modern Approaches to the Unseen Environment, Princeton Univ. Press, Princeton, N.J., pp. 512-543, 1950.

10. Smith, G.C. and Mauch, H.A. "Summary Report on the Development of a

Reading Machine for the Blind", Bulletin of Prosthetics Research, BPR 10-12, 243-271, Fall 1969.

11. Licklider, J.C.R. "The influence of interaural phase relations upon the masking of speech by white noise" J. Acoust. Soc. Am., 20, 150-159, 1948.

12. Hirsh, I.J. "The influence of interaural phase on interaural summation and inhibition" J. Acoust. Soc. Am., 20, 536-544, 1948.

13. Schubert, E.D. "Some preliminary experiments on binaural time delay and intelligibility" J. Acoust. Soc. Am., 28, 895-901, 1956.

14. Feldmann, H. "The role of interaural intensity differences and time delay for signal detection in noise" International Audiology, Vol. IV, No. 2, 29-34, 1965.

15. Levitt, H. and Ribiner, L.R. "Binaural release from masking for speech and gain in intelligibility" J. Acoust. Soc. Am., 42, 601-608, 1967.

16. Carhart, Raymond; Tillman, T.W. and Johnson, K.R. "Release of masking for speech through interaural time delay" J. Acoust. Soc. Am., 42, No. 1, 124-138, July 1967.

17. Carhart, Raymond; Tillman, T.W. and Greetis, E.S. "Release from multiple maskers: effects of interaural time disparities" ibid. 45, No. 2, 411-418, Feb. 1969.

18. Carhart, Raymond; Tillman, T.W. and Dallos, P.J. "Unmasking for pure tones and spondees: interaural phase and time disparities" J. Speech Hear. Res., 11, No. 4, 722-734, Dec. 1968.

19. Dirks, D.D. and Wilson, R.H. "The effect of spatially separated sound sources on speech intelligibility" J. Speech Hear. Res., 12, No. 1, 5-38, March 1969.

20. Cramer, H.L. "The intelligibility of time-compressed speech" Proceed-

ings of the Louisville Conference on Time Compressed Speech, Oct. 1966.

21. Carroll, J.B. and Cramer, H.L. "The intelligibility of time-compressed speech final report" U.S. Dept. of Health, Education and Welfare, Bureau of Research, Final Report OE 7-31-0370-271, July 1968.

22. Fairbanks, G., Everitt, W.L. and Jaeger, J.P. "Method of time or frequency compression-expansion of speech", IEEE Trans. Audio, Vol. AU-2, 7-12, 1954.

23. Dolch, E.E. et al. "Famous stories for pleasure reading", Garrard Publishing Co., Champaign, Illinois, 1955.

24. "1965 Revised List of Phonetically Balanced Sentences (Harvard Sentences)" IEEE Trans. Audio and Electroacoustics, Vol. AU-17, No. 3, 239-246, Sept. 1969.

25. Miller, G.A. "Speech and Language" in Handbook of Experimental Psychology, S.S. Stevens (editor), p. 794, John Wiley, 1951.

26. Hoops, R.A. Speech Science: Acoustics in Speech, Charles C. Thomas (Publisher) U.S.A., p. 68, 1969.

27. Parmenter, C.E. and Treviño, S.N. "The Length of the Sounds of a Middle Westerner", American Speech, 10, 129-133, 1935.

28. Ingham, K.R. "Spelled speech as an output for computers and reading machines for the blind", IEEE Trans. Human Factors in Electronics, Vol. HFE-8, 175-181, Sept. 1967.

29. Fairbanks, G. and Kodman, F. "Word Intelligibility as a function of time compression", J. Acous. Soc. Am., 29, 636-641, May 1957.