DESIGN OF A SIMPLE READING MACHINE FOR THE BLIND

by

WILLIAM DESMOND RAMSAY
B. Eng., Carleton University, 1966

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 Commitee

Head of Department

Members of the Department
of Electrical Engineering

THE UNIVERSITY OF BRITISH COLUMBIA

October, 1968
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical Engineering

The University of British Columbia
Vancouver 8, Canada

Date Nov 15, 1966
ABSTRACT

A compact reading machine ("Lexiphone") has been designed and constructed to convert printed letters into a pattern of sounds. The machine reads by direct translation of vertical sections of the letters, according to a recently developed code. In this code, the "melody" produced is independent of the vertical position of the reading head; however the user is given an indication (mean pitch) of the vertical position to facilitate tracking along a line of print.

The discrete nature of the direct translation process limits the theoretically possible reading rates. Tests with artificially generated codes were performed to investigate this limit, and it is expected that the limit will be above that for Morse Code--60 to 70 words per minute. This would be adequate for practical use.

Tests performed at Haskins Laboratories predicted similar performance for other machines, such as the optophone. However, practical users of the "Battelle Optophone", the most refined version of the optophone, attained only 25 words per minute (on Grade I reading material) after an extensive course. It is suggested that this was due to the difficulty in the earlier machines of producing repeatable versions of the code.

Code sounds from the present Lexiphone prototype were found to be very consistent and repeatable, and should allow the predicted reading rates to be approached. Practical reading results with the machine are presented. At the time of writing, a subject training with the machine is reading two-page passages of Grade III material at 30 words per minute, and her performance is still improving.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>viii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.2 The Problem</td>
<td>1</td>
</tr>
<tr>
<td>1.2.1 History</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2 Reading Machine Codes</td>
<td>5</td>
</tr>
<tr>
<td>1.2.2.1 General Considerations</td>
<td>5</td>
</tr>
<tr>
<td>1.2.2.2 The Lexiphone Code</td>
<td>7</td>
</tr>
<tr>
<td>2. THE HARDWARE OF THE LEXIPHONE</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Organization of the Machine</td>
<td>9</td>
</tr>
<tr>
<td>2.2 The Input Section</td>
<td>10</td>
</tr>
<tr>
<td>2.2.1 The Array</td>
<td>11</td>
</tr>
<tr>
<td>2.2.2 The Schmitt Triggers</td>
<td>14</td>
</tr>
<tr>
<td>2.2.3 The Mechanical Page Scanner</td>
<td>17</td>
</tr>
<tr>
<td>2.2.4 The Imaging System</td>
<td>17</td>
</tr>
<tr>
<td>2.2.5 Page Illumination</td>
<td>18</td>
</tr>
<tr>
<td>2.3 The Digital Section</td>
<td>19</td>
</tr>
<tr>
<td>2.3.1 Principle of f-Computation</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2 Digital Section Organization</td>
<td>21</td>
</tr>
<tr>
<td>2.3.3 Digital Section Circuit Details</td>
<td>26</td>
</tr>
<tr>
<td>2.3.3.1 Clock and Timer</td>
<td>28</td>
</tr>
<tr>
<td>2.3.3.2 Scanning Circuit</td>
<td>29</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>2.3.3.3</td>
<td>Edge and &quot;Count By&quot;</td>
</tr>
<tr>
<td>2.3.3.4</td>
<td>Counting Control</td>
</tr>
<tr>
<td>2.3.3.5</td>
<td>f-Counter</td>
</tr>
<tr>
<td>2.3.3.6</td>
<td>The Output Buffers</td>
</tr>
<tr>
<td>2.3.3.7</td>
<td>Error Protection</td>
</tr>
<tr>
<td>2.3.3.8</td>
<td>Top Edge Circuit</td>
</tr>
<tr>
<td>2.3.3.9</td>
<td>Intensity Modulation</td>
</tr>
<tr>
<td>2.3.3.10</td>
<td>D/A Converter</td>
</tr>
<tr>
<td>2.4</td>
<td>The Analog Section</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Exponential Function Generator</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Voltage Controlled Generator</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Amplitude Modulation</td>
</tr>
<tr>
<td>3.</td>
<td>PERFORMANCE CHECKS</td>
</tr>
<tr>
<td>3.1</td>
<td>Machine Tests</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Upper Outline</td>
</tr>
<tr>
<td>3.2</td>
<td>Code Tests</td>
</tr>
<tr>
<td>3.3</td>
<td>Reading Results with the Machine</td>
</tr>
<tr>
<td>3.4</td>
<td>Continuing Work</td>
</tr>
<tr>
<td>4.</td>
<td>SUMMARY AND CONCLUSIONS,</td>
</tr>
<tr>
<td>4.1</td>
<td>Remaining Questions</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1-1</td>
<td>Production of the Lexiphone Code</td>
</tr>
<tr>
<td>2-1</td>
<td>Block Diagram of the Lexiphone</td>
</tr>
<tr>
<td>2-2</td>
<td>Effects of Alignment Change</td>
</tr>
<tr>
<td>2-3</td>
<td>Testing with Prepared Films</td>
</tr>
<tr>
<td>2-4</td>
<td>Type Face Measurement</td>
</tr>
<tr>
<td>2-5</td>
<td>Lexiphone Prototype</td>
</tr>
<tr>
<td>2-6</td>
<td>Events During a Typical Scan</td>
</tr>
<tr>
<td>2-7</td>
<td>Functional Diagram of the Digital Section</td>
</tr>
<tr>
<td>2-8</td>
<td>Timer Outputs</td>
</tr>
<tr>
<td>2-9</td>
<td>Schmitt Triggers</td>
</tr>
<tr>
<td>2-10</td>
<td>Clock and Timer</td>
</tr>
<tr>
<td>2-11</td>
<td>Scanning Circuit</td>
</tr>
<tr>
<td>2-12</td>
<td>Edge and &quot;Count By&quot;</td>
</tr>
<tr>
<td>2-13</td>
<td>Counting Control</td>
</tr>
<tr>
<td>2-14</td>
<td>f-Counter</td>
</tr>
<tr>
<td>2-15</td>
<td>Conditions for Loading Output Buffer</td>
</tr>
<tr>
<td>2-16</td>
<td>Production of Top Edge Pulse</td>
</tr>
<tr>
<td>2-17</td>
<td>Typical f-Function for the &quot;I&quot;</td>
</tr>
<tr>
<td>2-18</td>
<td>Black Pulse Counting</td>
</tr>
<tr>
<td>2-19</td>
<td>D/A Converter</td>
</tr>
<tr>
<td>2-20</td>
<td>Analog Section</td>
</tr>
<tr>
<td>3-1</td>
<td>Raw Signal from a Photodiode</td>
</tr>
<tr>
<td>3-2</td>
<td>Effect of the Schmitt Triggers</td>
</tr>
<tr>
<td>3-3</td>
<td>Print Viewed by the Lexiphone</td>
</tr>
<tr>
<td>3-4</td>
<td>Top Edge Recognition</td>
</tr>
</tbody>
</table>
3-5  Spurious Pulses................................. 45
3-6  Unavoidable "Errors"............................. 46
3-7  Tests of the f-Computation...................... 48
3-8  Measured f-Function - Test #1.................. 49
3-9  Measured f-Function - Test #2.................. 49
3-10 Measured f-Function - Test #3.................. 50
3-11 Sample of Practical f-Function................. 50
3-12 Comparison of Codes............................ 53
3-13 Reading Results with the Lexiphone............. 56
3-14 Battelle Reading Speeds......................... 57
LIST OF TABLES

Table | Page
-----|------
3-1   | Transliterations to Produce Wuhzi.. 52
3-2   | Wuhzi Equivalents of Words in the Eight Word Test.. 52
ACKNOWLEDGMENT

Grateful acknowledgment is made to the many people who assisted during the course of the project.

Particular thanks is due to my supervisor, Dr. M. P. Beddoes, for his encouragement and advice.

I would like to thank Professor F. K. Bowers for his careful reading and correction of the manuscript.

Blind subjects Peggy Spencer and Linda Jentsch deserve recognition for the many hours they spent working with the Lexiphone. Also, thanks go to the volunteers from the department who subjected themselves to the Lexiphone and other code sounds.

I would also like to thank Peter Dewdney, Willis Martin and Hans Paul for proofreading the thesis, Grant Murray for his work building equipment, and Avis Hopkins for the typing.
1. INTRODUCTION

This thesis deals with the design, construction and testing of the Lexiphone, a relatively simple reading machine for the blind. The machine reads the printed page and produces a musical code as output. The musical code was proposed by Beddoes (1) and tested on computer simulation. The present machine produces a repeatable version of this code and has been quite dependable. The thesis also includes results of code tests that were performed in an attempt to explain the failings of previous machines.

1.2 The Problem

The problem of enabling the blind to read the standard printed page is a very old one, but is still essentially unsolved. The blind person who wants to read the local news, notes from friends, bulletins, or any of much similar material not yet translated into braille has no choice other than to use a sighted reader. This is a tremendous inconvenience, and the lack of independence and privacy is very distasteful to many blind people. Any readily available device that would allow even the reading of personal correspondence would be a great improvement indeed.

Almost since it was first technically possible, attempts have been made by various workers to make such a simple machine. Although many machines have been constructed, reading with them has been such a tedious and painfully slow process that none has yet gained acceptance as a useful aid.

The Lexiphone belongs to the group known as "direct translation machines". A "direct translation machine" considers only the
patterns of black and white presently under its scanner, and using some transformation, produces a sound that represents what it sees. The direct translation machine thus produces an instant-by-instant translation of print patterns into sound. It is then the task of the user to find out which sound patterns correspond to particular letters and words, the same way as a sighted reader must learn which patterns of ink correspond to given letters and words. Such a machine can obviously be used for any reading task. The problem is whether or not the sounds it makes can be easily and quickly interpreted.

Due to their potential low cost and portability, direct translation machines have been the ones most widely considered in study of the problem of providing a personal machine. Previous machines have been sufficiently limited by technology that designers concentrated on making a sound from ink, hoping that it would produce a suitable code when print was scanned. In the Lexiphone, the philosophy has been that one should feel free to find a good code, confident that it can be instrumented. Another problem typical of older machines was poor resolution which caused the code to be very dependent on alignment. This resulted in machines that were difficult to operate and in code sounds that lacked repeatability. Some of the older machines are described briefly in the next section.

1.2.1 History

A complete historical review has been given by Caple. The principal machines built to date are as follows:

1914: White Sounding Optophone - The work of Fournier d'Albe, it had eight pulsed areas of light on the page and gave a tone
when any area was reflected to the cell.\(^{(3)}\)

1920: **Black Sounding Optophone** - Barr and Stroud, a British firm, put a five tone optophone in a bridge arrangement so that the sound was heard only when print was under the scanner.\(^{(11)}\) Bridge balance was poor.

1928: **Visagraph** - R.E. Naumburg proposed a machine that embossed a raised version of the letter on aluminum foil.\(^{(12)}\) It turned out to be harder to read than Optophone code.

During World War II, the U.S. Veterans Administration formed the Committee on Sensory Devices. They commissioned Haskins Laboratories and RCA to look into the problem of reading machines.

1944-49: **RCA A-2 Reader** - A swept audio oscillator coupled to a swept spot of light and keyed on when the spot was over print.\(^{(13)}\) Produces an optophone-like sound but has some noise due to keying.

**RCA Recognition Machine** - Output from a flying-spot scanner was modulated depending on the total amount of light reflected from the page as the light slit was passed over it.\(^{(4)}\) Sound was turned off when no letter was present. Reading was very slow.

(In addition to the work on the FM scan machine, the Haskins Laboratories evaluated other existing machines and tested different codes. This is discussed later.)

1952: **The Argyle Reader** - Argyle of British Columbia made a machine which had a spot of light scanning the letter vertically at 200 Hz, and the reflected light was picked up by a cell and
amplified. The result was an erratic buzz at harmonics of 200 Hz that could be read no better than the optophone.

1957 onward: Since 1957, the Battelle Memorial Institute, under the auspices of the Veterans Administration, modernised the original optophone, using a photocell for each tone rather than the pulsed areas of light. They built several of these improved machines and developed extensive training programs. In spite of this, the optophone has not been successful.

Visotoner, Visotactor, Cognodictor - Mauch Laboratories in Ohio have developed optophone-like machines with audible output (Visotoner) and with tactile output (Visotactor). They have also done some work on producing spelled speech output (Cognodictor). The Cognodictor is currently still in the experimental state. Such a machine might provide an output which would be simple to learn.

Tactile Machine - Linvill and Bliss have produced a machine which gives a tactile display of the letter on an 8 x 5 array of vibrating reeds.

Note that of these machines, those producing an audible code output are still similar to the original optophone in at least two important aspects. They produce generally a chord-like output, and are notably lacking in resolution. It is suggested that in these factors can be found the principal reasons for lack of success of the early machines.

Although all the machines made it possible to decipher a word eventually, with enough training and enough time, no machine was sufficiently easy to use. In the Lexiphone, some weaknesses of
previous machines have been removed. It is hoped that this work will help answer the question of whether the direct translation machine deserves further consideration.

1.2.2 Reading Machine Code

The machine described in this thesis has been designed to produce the Lexiphone code as proposed by Beddoes. The code has been tested in computer simulations at MIT and also here at UBC. It differs from less sophisticated codes in some significant respects. The melody produced is independent of the vertical alignment of the scanner and consists of only one note at a time. Chords are never played. In addition, there is amplitude modulation determined by a separate criterion.

1.2.2.1 General Considerations

Several workers, notably Nye, Cooper, and Zahl, have examined in detail the problem of aural outputs. The conclusion seems to be that no artificial letter-by-letter code has any hope of competing with the natural speech code. However, there seems to be no fundamental reason why the rate of reading of such a code should not be fast enough to make it of use.

Cooper and Zahl decided that the limiting factor was the blending of the code elements into a blurr or a buzz, and used the reception of International Morse code as a case in point. Much the same idea was expressed by Caple when he related the reading speed to the maximum channel capacity of the subject in bits/sec.

In speech, the rates vary from about 120 wpm, typical of radio announcing, to between 150 and 200 wpm for normal conversation.
In the case of International Morse, the only encoding is in amplitude (on-off keying). A letter has from one to four elements, the average value being three elements per letter. If we assume that the speed at which elements start to blend into a buzz is 20 Hz, it gives a reading rate of 60-70 wpm as an absolute maximum for the Morse. Speeds of this order have been recorded in exceptionally gifted and experienced operators. Commercial operators handle code comfortably and dependably at speeds of about 35 wpm. The DOT requires a speed of 15 wpm for the Advanced Amateur licence. Thus the first plateau in reading speed, attained after a fairly rigorous formal course, is about 20% of the absolute maximum, and the useful rate for experienced operators is about one half the maximum value.

Cooper and Zahl\(^4\) extended this experience to reading machine codes. They conclude that there are about the same number of elements per letter in the optophone code as in the Morse, and as a result assume that we cannot expect much more by way of performance with the optophone than with the Morse code. The Lexiphone code is modulated in both amplitude and frequency. The frequency is modulated with a composite upper outline of the letters, and the amplitude is modulated by making the sound louder when the scanning slit is over a vertical line, or "riser". In contrast to the Morse where only two elements, the dot and the dash, are used to code the letters, the Lexiphone code has a host of frequencies to choose from, and, in addition, has amplitude modulation. The differences in modulating schemes should produce very different problems in assimilation. Results of some tests with this and other codes are presented later in the thesis. When letters are presented slowly in Lexiphone code, a wealth of detail is perceived. This detail may be used by an
operator in identifying an isolated letter. The letters are, however, overspecified. Thus, when heard at a faster rate, certain mean values may be sufficient for identification. When an entire word is listened to at high speed, what is perceived is a characteristic melody due to the dominant high and low frequencies in its component letters, and a characteristic rhythm due to the riser thumps which occur only in certain letters. A striking demonstration of the low number of psychologically perceived elements in a word is the fact that a user finds that he himself starts to mimic Lexiphone coded words. It is due to this formation of "gestalten" (single, whole images formed from discrete elements) that decoding can proceed at a rate above the Cooper-Zahl buzz frequency.

12.2.2 The Lexiphone Code Transformation

As mentioned above, the Lexiphone code is frequency modulated by a composite upper outline of the letters and amplitude modulated by the presence of a vertical line or "riser" (e.g., "1", "l", etc.)...
"d", "h", "k", etc.). The actual transformation, proposed by Beddoes(1) is shown in Fig. 1-1. The frequency values range from 200 Hz to a maximum of 3000 Hz. The amplitude can have only two values, loud or soft. The composite upper outline, or f-function, is determined only by the position of the upper edges of the print information. The transformation is thus independent of letter thickness.

To form the f-function, the distance from the bottom of the scanning slit to the highest upper edge \( y_m \) is added to the sum of the distances down to the other upper edges from the topmost outline. \( \Delta y \)'s in Fig. 1-1) The f-function at any point in the letter represents the musical note that will be heard when the scanner is at that point in the letter. The relation of output frequency to f-function has been arranged (Section 3.4) so that if the letter is shifted up (or down) in the scanning slit, it just moves the f-function up or down on the musical staff, and exactly the same melody is heard, but in a different key. The changing key provides an immediate clue to the alignment of the scanner, while the constant melody means that comprehension is not affected. By merely listening to the key of the melody of the print, the operator knows immediately if he is drifting up or down on the line. Blind subjects using the present Lexiphone prototype depend on this pitch clue to determine the correct vertical alignment.

The intensity modulation due to the risers provides the rhythmic pattern, which, in conjunction with mean letter frequencies, are helpful in formation of gestalten, the key to rapid assimilation of code sounds.
2. THE HARDWARE OF THE LEXIPHONE

2.1 Organization of the Machine

The organization of the machine can be seen from the block diagram of Fig. 2-1. The image of the print is focused on to the photocell array. This gives rise to 54 analog currents, each varying in accordance with the instantaneous illumination of that particular cell. The analog photocurrents are then passed into Schmitt triggers which give the digital signals for the next section.

The heart of the Lexiphone is the block labelled "f-computer". This unit takes the signals from the Schmitt triggers and calculates the f-function according to the Beddoes transformation (Fig. 1-1). Parallel to the f-computer we have the riser detector which decides when the array is over a vertical riser and produces the required amplitude modulating signal.

The D/A converter puts the f-function into analog form convenient for modulating the audio generator.

The frequency modulating signal is produced by passing the analog f-function through an exponential element. The exponential transfer function assures that fixed voltage intervals of the f-function produce fixed musical intervals in the audible output. The amplitude modulation is three valued, i.e., OFF when no letters are seen, LOUD when a riser is indicated, and SOFT otherwise.

A more detailed description of each section follows.
2.2 The Input Section

One of the biggest problems in instrumentation of a usable reading machine is translation of a normal printed page into a signal suitable for processing by a code generating unit. A method used in previous reading machines, and the one chosen for this prototype of the Lexiphone, is to focus the image of the print to be read on to an array of photocells. While simple in principle, this poses many practical problems.
2.2.1 The Array

The array used in the machine is the Fairchild FPA-500 linear integrated photodiode array. It has two columns of photodiodes, each column having 63 diodes on 12 mil centres. Fifty-four cells of one column are used in the Lexiphone. Each diode is 35 mil\(^2\) in effective area, with a 4 x 6 mil region of uniform light sensitivity. An elemental diode of the array has a dark current (at room temperature) of less than 1 na, and the photocurrent rises linearly with irradiation to about 750 na at 30 mw/cm\(^2\).

Each cell of the array was measured by Fairchild before delivery at 5 mw/cm\(^2\) irradiation* and gave a typical photocurrent of 150 na. Actually, the light levels required for this amount of irradiation are rather impractical, and it was necessary to operate the array in the Lexiphone at much lower photocurrents.

Previous machines had used only a small number of cells, (typically 10) and, as a result, resolution was not good and the code was very dependent on alignment. In the case of the optophone, there was a good reason for not going to a larger number of cells. The optophone code consists of a tone for each cell, and becomes prohibitively confusing as the number of cells is made large. The Lexiphone code, on the other hand, depends only on the measurements of the letter itself, and the purpose of the scanner is to get an accurate measurement of the letter. The pictures of Figure 2-2

* As the photodiode array is quite sensitive to wavelengths in the long visible and infrared regions of the spectrum, it is more suitable to measure the sensitivity in terms of the total irradiation (measured in, say, mw/cm\(^2\)) than the illumination, which is visible only. For a given incandescent source, the fraction of the total power radiated that falls in the visible is a function of the colour temperature. For example, one mw/cm\(^2\) of irradiation from an unfiltered tungsten source operating at 2870\(^\circ\)K constitutes about 20 ft.-candles of illumination.
Fig. 2-2 Effects of alignment change on high and low resolution scanners. (a) Source material of 3 alignments. (b) Picked up by high resolution scanner. (c) Picked up by low resolution scanner.
provide a simple demonstration of one of the benefits of a high resolution scanner. They show "G" as seen by the present 54 cell scanner; and by an array of only 6 cells typical of earlier machines. As we see, the low resolution scanner sees quite a different letter when its vertical alignment is changed (The fact that it does not really look like a "G" would not matter if it could be made to come out the same for all alignments).

The nature of the 54 cell array has another advantage in connection with the signal from an individual cell. An individual cell of a high resolution array covers a much smaller area of the letter and as a result gives a sharper transition when moving from a light to a dark region. Later in the thesis is shown the waveform from a cell, taken as the machine scans a row of vertical lines. The well defined white-to-black dip is very evident (Fig. 3-1).

To reduce the number of variables in the testing of preliminary circuitry, test material was prepared on high contrast 35 mm film and pulled directly across the surface of the array (Fig. 2-3). This method provides a nearly ideal optical signal to the array. It also puts all the important factors for testing, such as intensity of illumination, size, position, and style of type, completely under control.

![Diagram](image)

Fig. 2-3 Testing with prepared films.
The size of a given type face is measured from the top of a riser to the bottom of a descender as depicted in Fig. 2-4. This measurement is generally given in "points", where 72 points equals approximately one inch. The type on the typewriter used for most of the test material measured 0.125" (9 point), and required a magnification of four to fill 80% of the FPA-500's 640 mil active area. The magnification should never be such that the type face fills the entire array, as this would leave no latitude in alignment. Another array has been purchased for use in future machines. It is the FPA-203, which has only 63 diodes on 4 mil centres. It has only one quarter the sensitivity of the FPA-500, but the 4 mil centres will reduce magnification requirements by a factor of three, which more than makes up for the decrease in sensitivity.

![Type size measurement diagram](image)

Fig. 2-4 Type Size is measured in "points" from the top of riser to bottom of descender. (72 points = 0.9962")

2.2.2 **The Schmitt Triggers** (Fig. 2-9)

Essentially, the function of the rest of the machine is to take the raw signals from the photocell array and produce the musical
code that would be arrived at by measurement of the print. The Schmitt triggers accept the cell outputs and produce a logical 1 (ground) on cells that have less than some certain value of light current, chosen to represent the highest value likely to be obtained when the cell is over ink. There are many instances, say just on an edge or over a thin or hazy line, that the current can be in an indeterminate region between light and dark. The Schmitt triggers eliminate this uncertainty. (Fig. 3-1 and 3-2).

Fig. 2-9 Schmitt Triggers

Since the photodiode is such a high-impedance device, an FET input has been used in the trigger circuit. This unit has a hysteresis of 0.2v, which is about 1/5 of the swing from black ink to the white page. An individual setting pot has been included on each trigger. This compensates for variations in FET characteristics; and
Fig. 2-5  (a) present prototype
(b) compact electronics for further machines
also for non-uniformities in page illumination.

2.2.3 The Mechanical Page Scanner (Fig. 2-5)

In this prototype, instead of moving the photocells over the page, the page is moved under the photocells. The same mechanism was used by Caple in his Lexiphone. Two modifications were made: the addition of the adjustment screws for alignment of the table during setup, and the addition of an instant-stop lever so the operator may control the table himself. This method is, of course, relatively unsatisfactory for a functional reading machine, but it does serve as a convenient test device. The actual product would probably use a hand-held reading probe such as the type used by Mauch in his Visotoner and Visotactor. This probe is moved across the page by the operator and will give him complete control of reading speed and alignment.

A system employing fiber optics has been proposed that may simplify the reading head (see Section 4.1).

2.2.4 The Imaging System

At present, a camera lens (Elgeet Cine Navitar 12 mm, F1.2) is used to produce the desired image. Such a lens is far too bulky to include in a hand-held probe. It was used merely because it is versatile and was available for use at the time. The lens chosen should have a short focal length to keep the physical size of the system down. It should have as wide an aperture as practical (the F1.2 of this lens is quite good) to reduce illumination requirements. One immediately apparent drawback of a wide lens opening was the very shallow depth of field. Even irregularities in the page caused a loss of focus and it became necessary to put a glass plate
over the test paper. A hand-held probe would provide direct contact with the page and focus would not be a problem.

2.2.5 Page Illumination

As discussed in Section 2.2.4, the distinct disadvantage of a high-resolution scanner is the small amount of light that any one cell of the array is able to gather. To get useful amounts of photocurrent from the array, very high intensities are required on the page. At present, illumination is supplied by a Carl Zeiss microscope illuminator with the diffusion screen removed so that the special 15 watt bulb's filament is focused on the page. This particular bulb (6v 15w Zeiss #380177) has a flat wound filament that focuses nicely into a bright rectangle (about 0.2" x 0.4") which is convenient for illuminating a letter. This particular arrangement provides an illumination on the page of about 12,000 foot candles, which is more than bright sunlight, but only a small part of the light shining on the page makes its way to the photodiode array. Of the light scattered by the page, only a small amount will actually be gathered by the lens, and this is further reduced by the magnification.

\[ I_{\text{page}} = k (Fm)^2 I_{\text{array}} \]  

(3.1)

where \( F \) is the lens opening (Fl.2 here), \( m \) the desired magnification, and \( k \) a constant determined by the reflectivity of the page and the manner in which the light is scattered by the page.

With the above in mind, it is not surprising that the photocurrent from a diode of the array is only 60 na over a white page, indicating an irradiation of some 2.5 mw/cm² (about 50 ft. candles with the lamp used). Even though this value of light is in the lower
end of the array's useful region, it is not felt that it would be practical to attempt to boost the illumination on the page any further, at least certainly not by increasing the size of the lamp. Since the array only "looks at" a thin vertical section, the effective illumination might be increased by using a cylindrical lens to focus the light into a line. With the present machine such measures were not warranted.

2.3 The Digital Section

The digital section must take the signals from the Schmitt triggers of the input section and produce the f-function and riser signals needed by the final analog section. The digital section constitutes the bulk of the machine.

2.3.1 Principle of f-Computation

As noted in the section describing the code, the frequency of the audible output is:

$$\omega = ke^f$$  \hspace{1cm} (3.2)

where $f$ is a function calculated from:

$$f = y_m + \sum \Delta y$$  \hspace{1cm} (3.3)

During operation of the machine, the cells of the array are scanned from top to bottom, each completed scan producing a value for the $f$-function. The frequency of top to bottom scanning is approximately 10 KHz, yielding a virtually continuously changing $f$-function as the array is moved across a letter. During any one scan, each cell is interrogated in turn until the first white to black transition (top edge) is encountered. Each further interro-
A- Start of scan. "Count by" register contains 4 - the edge count plus one. No counting takes place.

B- 1st edge encountered. "Count by" is reduced by one, bringing it to the correct value of three. Counting by three commences.

C- 2nd edge encountered. "Count by" reduced to two. Counting continues by two.

D- 3rd edge encountered. Counting continues by one.

E- End of scan. Check is made that counting is by one and the value of count shifted into output buffer.

\[ f = y_m + \sum_{i=1}^{M-1} \Delta y_i \]

where \( M \) = number of black areas in scanning slits field.

Fig. 2-6 Events During a Typical Scan
gation, from this time until the next top edge is detected, adds to a counter a number \( N \), equal to the number of top edges presently under the slit (\( N \) was determined by the previous scan). From the second to the third top edges, \( N-1 \) is added to the counter for each cell sampled. This sequence continues to the bottom of the array. At the bottom, a check is made to see if the counting is now proceeding by one. If it is, it means that the number of top edges, \( N \) (borrowed from the previous scan), was, in fact, the correct value to use for this scan. This being the case, the total count will be the desired \( f \), and may be shifted out.

That this procedure actually does produce the function

\[
f = y_m + \sum \Delta y
\]  

(3.3)

can be seen from the example of Fig. 2-6. From B to C, each cell spacing counts three units, from C to D, two units, and from D to the end (E), one unit. This forms the same total as adding the three vectors \( y_m, \Delta y_1 \), and \( \Delta y_2 \).

2.3.2 Digital Section Organization

Reference to the block diagram of Fig. 2-7 shows how the above scheme has actually been realized using logic circuitry. The machine operates on a 64-step cycle, during which time the 54 cells are scanned and the \( f \)-function calculated. The timing of a cycle is detailed in Figures 2-6 and 2-8. The function is accumulated in the \( f \)-counter, which is simply an 8-bit up counter which can be toggled on the 4's, 2's, or 1's bit. Once enabled, the counting control circuit directs pulses to the appropriate inputs depending on the contents of the "count by" register. For example, if the machine
**Fig. 2-7 Functional Diagram of the Digital Section**
is counting by three, a pulse is first applied to the 2's and then to the 1's bit during each clock period.

On Fig. 2-7 it is noticed that the edge counter is set to 1 at the start of each cycle, rather than being cleared to zero. The reason for this is that the first top edge encountered in a scan should start the counting by the number of top edges under the slit. However, since each top edge pulse also reduces the "count by" by one, it is necessary to have the initial content one higher than the desired starting count. This has been accomplished by having each scan start with the edge counter already at one.

The master timer controls the sequence of all operations of the machine. It generates three control pulses, $P_1$, $P_2$, $P_3$, and an analog voltage proportional to the cycle step number. It is in turn controlled by the clock.
The Letter Present (LP) flip-flop is set by the black pulse, $P_b$, and hence indicates the presence of black under the scanner. The Edge Present (EP) flip-flop is set by the top edge pulse, $P_t$, and hence is 1 when a white-to-black transition (top edge) has been detected. The muting FF is used to silence the output of the machine during the space between letters (The letter present may not be used for this purpose directly, as it is reset at the start of each scan).

The sequence of operations is as follows:

$P_1$:
- Loads output buffer with contents of the $f$-counter if the letter present and edge present are in the same state and if the "count by" is at one.
- Loads muting FF from the LP.
- Stores output of a circuit counting the number of cells black (used for riser detection).

$P_2$:
- Clears the $f$-counter to zero.
- Resets EP FF (this disables the counting).
- Resets LP FF.
- Resets the "black count" FF.
- Loads edge count into the "count by" register.

$P_3$:
- Loads intensity FF.
- Sets edge count to one.
- Starts new scan.

A feature not yet discussed is the intensity FF. It is used to make the intensity of the code output greater when only one edge is present and more than eight cells are indicating black. At
time $P_2$, the "count by" is loaded with the contents of the edge counter which will be one greater than the number of top edges under the slit. At time $P_3$, then, the circuit looks at the $n=2$ line in the counting control and if it is HI, indicating only one edge under the slit, then the intensity FF is set to make the sound loud. If, in addition, the output of the black counting circuit is at ground, indicating more than eight cells black, the sound is intensified.

The time from $P_3$ to the end of the scan is occupied by the Black Pulses, $P_b$, produced when cells over black are sampled, and the top edge pulses, $P_t$, produced by the required white-black sequence in $P_b$.

$P_b$:  
- Sets the Letter Present FF.  
- Removes the blanking level from the CRT.  
- All $P_b$ go to the top edge decision circuit where the sequence is examined and $P_t$ produced at the right time.

$P_t$:  
- Sets the Edge Present FF (this enables counting).  
- Reduces "count by" by one.  
- Adds one to the edge count.

It was stated before that there are certain conditions which must be met before the value of "f" in the counter can be transferred to the output buffer and hence used for the code. The proviso that the edge count must be equal to one has already been discussed and is due to the fact that the top edge count used is the one from the previous scan, and we must protect against a wrong value when the number of top edges changes. In some cases, very thin lines,
or light parts of print, a condition may exist where black is seen by the scanner, but no top edge is indicated. In this case the transfer of "f" will be inhibited by another circuit, as the value is obviously wrong. The circuit would also protect us in case of a $P_t$ without a $P_b$, although this will never occur in practice.

On good print, these precautions produce virtually no change in the sound of the final output. Before the machine was in its final form, the muting was done by looking at the contents of the output buffer, and the output was muted when this was zero. In the case of a top edge dropout ($P_b$ with no $P_t$) this caused the output to be muted completely during the dropout. The result was a very raspy sound on print where such dropouts were a problem. With the muting done by looking for ink, the output is not muted in the case of a dropout, and the problem is not so acute. Nevertheless, the protection is very simple and has been left in. Similarly, the error due to a change in top edge number is slight and occurs only on that scan in which the number changes, but, again, protection is so simple it has been left in.

The analog section has been shown as a box in Figure 2-7. It is detailed later. It produces the desired audio output from the f-function, the muting, and the intensity levels.

2.3.3 Digital Section Circuit Details

Motorola MC700P series integrated circuits have been used extensively in the digital section. Some Fairchild 900 series were used in the early stages of the machine but supply became difficult, and machines presently being built use Motorola throughout.
to the desired combination on gates giving $P_1$, $P_2$, $P_3$.

$R = 1.2K$

$2R$ detail

$3.3K$ $10K$

Fig. 2-10 Clock and Timer
2.3.3.1 **Clock and Timer** (Fig. 2-10)

The clock is a free-running multivibrator which gives square wave output at about 600 KHz. The timing is done by the 6-bit up counter, giving a 64-step cycle. The counter was convenient for this purpose, as the D/A ladder could be connected to give

![Clock and Timer Diagram]

**Fig. 2-11 Scanning Circuit**
an output voltage proportional to the step in the cycle. With this analog ramp on the vertical deflection of a 'scope and the black pulse connected to modulate the beam intensity, we can get a good display of what is being seen by the scanner. This turned out to be a very valuable aid indeed in setting up the machine. If the array is not exactly vertical, or if there is poor focus or uneven lighting, this is easily detected by looking at the display of the print on the 'scope. In a final model of the machine, this feature would probably be omitted. (Figure 3-3 shows an example of the above 'scope display.)

2.3.3.2 Scanning Circuit

The purpose of the scanning circuit is to sample each cell of the array in turn and to produce a pulse on the cells that are indicating black. It is essentially a shift register and a set of gates connected to the 54 lines from the input section. The scanner is detailed in Figure 2-11. On P₃, the scan is begun by starting a "one" out at the start of the shift register. The ground enable level is thus applied to each gate in turn and every time the corresponding trigger is indicating a black, a pulse is produced at the gate output. All the gate outputs are fed to a diode OR circuit.

The trigger levels of the Schmitt circuits in the input section are adjusted as follows: a white page is examined, but with reduced illumination level. The potentiometers are then set so that each of the circuits just indicates black. This gives a P_b signal that is an accurate reproduction of the black areas passing under the array. The 'scope displays obtained from the P_b signals (see test section, Figure 3-3) attest to the resolution of the array and the
cleanness of operation of the triggers. This is an important point in the present work, as it is felt that the failure of previous attempts to make direct translation machines has been, in large part, due to the poor resolution and poor repeatability of the scanning arrangement.

2.3.3.3 Edge Counter and "Count By" Register

This uses the circuit of Figure 2-12. \(P_3\) sets the edge count to one; \(P_t\) adds one to the count each time; \(P_2\) transfers the value of the edge count to the count by register; and \(P_t\) also decreases the "count by" by one each time.

![Diagram of Edge and "Count By"](image)

Fig. 2-12 Edge and "Count By"
2.3.3.4 Counting Control (Figure 2-13)

The counting control produces the appropriate pulses to the 4's, 2's, and 1's bits of the f-counter as the "count by" is changed. The string of monostables produces a sequence of three pulses during each clock period and they are directed to the appropriate counter inputs. It is important that the pulses be sequential so that any carry in the f-counter will have a chance to propagate before the next bit is added. The four gates decode the "count by" and the diode matrix then disables the appropriate counting bits. For example, if the count is to be "by two," the n=2 line from the gates will be HI, and the diodes will apply disable levels to the 1's and 4's bits.

The counting control also provides the necessary n=1 and n=2 outputs.

2.3.3.5 f-Counter (Figure 2-14)

The f-counter, as was mentioned before, is an 8-bit binary up-counter so wired that it can be toggled on the 4's, 2's, or 1's bit. Capacitive coupling is used so that the level on a FF output does not disable the gate used to OR the carry with the external input on that bit. In this circuit, any carry will propagate immediately, but there is a recovery time, set by the period of the monostables, during which time no further carry can propagate. It is therefore important that the recovery time of the f-counter be less than the time between successive pulses from the counting control. In the present machine, the clock period is 1800 ns. The length of each monostable in the counting control has been adjusted to approximately 400 ns, and the pulse length of the coupling monostables in the f-counter is 100 ns.
Fig. 2-13 Counting Control

Fig. 2-14 f-Counter
2.3.3.6 **The Output Buffer**

The output buffer stores the value of the \( f \)-function during a scan until the new value is ready. It is composed of 8 JK flip-flops connected to the \( f \)-counter and loaded by pulsing their toggles.

2.3.3.7 **Error Protection**

The protection against loading the output buffer if \( n \neq 1 \) or if the LP and EP are not in the same state, is shown in Figure 2.15. This will prevent the loading of an erroneous value of \( f \) in the case of a top edge dropout or when the number of top edges changes.

Fig. 2-15 Conditions for Loading Output Buffer

2.3.3.8 **Top Edge Circuit**

The top edge circuit deserves a special discussion, as it is in this region, as well as in the area of deciding which criterion to use for the intensity modulation, that a good deal of experimen-
tation and testing was necessary. In the case of the top edge, the problem is essentially this: we must make a decision as to which sequence of black-white to produce a top edge pulse on. If we decide just one black is automatically an edge, then we get each little bit of dirt and also get an edge indicated if there is a solitary white in the middle of a huge island of black. On the other hand, if the decision is made to look for a large number of black cells, then we will have trouble on legitimate horizontal lines of thin print. A good general method, and the one that was eventually decided upon, to produce a top edge pulse is to feed the P_b's into a short b-bit shift register where we can examine the last b cells sampled and decide when a top edge has occurred (see Figure 2-16). The final model used b=5 and a code of 11000, which means we must have two white cells and two black cells, separated by a "don't care" cell. Blind subjects presently using the machine find the code quite consistent on a variety of prints.

When the array is passing off the edge of a vertical line, there is a situation in which the state of the cells is more or less random. This will occur for an instant on the sides of vertical lines no matter how clear they are. If it were possible, the top edge code should also be one that helps prevent a huge number of top edges from being indicated when the cells are in such a condition. Actually, the duration of such bursts is so short that it is not important to suppress them. Figure 2-17 shows the f-function for the capital "I".
Fig. 2.16 Production of Top Edge Pulse

gates connected to give $P_T$ on desired $P_B$ sequence

Fig. 2.17 Typical $f$-function for the "I". Various features are worth noting here. Notice that the "I" was slightly tilted. The scanner first encountered the high top edge, then the two serifs, next the riser (same height as the first edge), then the two serifs again and finally the lower serif. Notice that the error signals as the scanner leaves the vertical edge are of very short duration.
2.3.3.9 The Intensity Modulation

The main use of intensity modulation is to provide additional clues to the identity of a letter or word, other than those provided by the frequency modulation alone. Without intensity modulation, some of the letters are just about impossible to tell apart. For example, the lower case "e" and the lower case "s" have almost identical f-functions. "M" and "Y" are identical in f-function.

The sound is made louder when the scanner is over a vertical line, or "riser", such as the riser in the letter "d", or the three risers in the letter "m". The machine indicates a riser whenever only one black island is present under the scanning slit and more than eight cells are black. The presence of a solitary island of black (indicated in the machine by only one top edge pulse) is almost sufficient to define a riser. The proviso that more than eight cells be black must be added to prevent indicating a "riser" on a single horizontal line.

![Fig. 2-18 Black Pulse Counting for Riser Detection](image)
Since the intensity FF is set by $P_3$, the information as to whether or not there was enough black present to indicate a riser must be available at this time. The method is to set a FF when the count in a 4-bit up counter passes 8 and to store this information for use in the next scan. (Figure 2-18)

Fig. 2-19 D/A Converter. (a) Weighted currents are switched to the operational amplifier by the diodes. (b) Detail of a bit switcher. $R = 470$ in all but MSB where it is 220 so that the entire current can flow with the available voltage difference. (c) Detail of a resistor of the D/A.
2.3.3.10 The D/A Converter (Fig. 2-19)

The D/A converter is the one described by George Austin.\(^{(7)}\)
In the converter, each bit of the 8-bit word can switch the appropriately weighted current into the input of the operational amplifier. The amplifier then sums all these currents and produces the word in the output buffer. In the Lexiphone, this analog voltage is the f-function itself, which we must use to produce the sounds of the Lexiphone code. This work is done by the analog section of the machine.

2.4 The Analog Section

In a musical scale, frequency ratios, not frequency intervals, are important. For example, on the piano keyboard, a given spacing of keys produces the same musical interval (octave, fifth, third, etc.), no matter where on the keyboard it is played. The Lexiphone is to be a musical machine. If we were to imagine the lowest value of the f-function possible, as sounding the bottom key on the keyboard, then we would expect as the f-function increased linearly in value, the note sounded would move linearly up the keyboard. Notice, however, what this means in terms of voltage to frequency translation. If each time the function increased by 2 volts, it moved us up the keyboard one octave, then the frequency is doubling for each 2 volt increase. This is an exponential relation. Regardless of the exact scale chosen, the requirement for a musical relation is the same - a certain increment in the f-function must always produce the same ratio in frequency. In the Lexiphone, the musical relation has been obtained by feeding the f-function to a circuit with an exponential transfer function and the output of this to a simple linear voltage controlled generator (VCG).
2.4.1 The Exponential Function Generator

The exponential transfer function of this circuit is obtained from the current-voltage characteristic of a diode. The voltage is fed to the diode, and the resultant, exponentially related current is converted to an appropriate voltage by the operational amplifier. The exponential element is provided mainly to give a position independence to the code, and any test of its quality will be more qualitative than quantitative. As such, the requirement is not an exacting one. The current voltage characteristic of a diode turned out to be more than adequate.

2.4.2 The Voltage Controlled Generator (VCG)

The VCG is really a modified UJT relaxation oscillator. The capacitor on the emitter of the UJT is charged from a transistor acting as a constant current source of adjustable value. Since the peak point voltage (at which the capacitor is discharged), remains constant, the period of oscillation will be linearly dependent on the charging current, and hence on the input voltage. A FF is used to produce square wave output. The pleasant oboe-like sound of the square wave is preferred by subjects.

2.4.3 Amplitude Modulation

The output of the reading machine can be loud, soft, or off. Both the on-off and the intensity switching functions are performed by FET switches.
3. PERFORMANCE CHECKS

3.1 Machine Tests

The first point of interest in a check is the actual photocurrent from a cell of the array. Figure 3-1 is the raw photocell output as the machine scans a row of brackets. This test was found to be very good for focusing the machine. The Lexiphone is set scanning the line and the focus adjusted to give the deepest dips as brackets are passed over. Notice that even on the raw signal, the white to black transition is very definite. Figure 3-2 is the signal taken after the Schmitt triggers. All the gray range has been eliminated and we have a definite white, or a definite black. It is these high-quality digital signals that are used to get the black pulse and finally the f-function.

The waveforms before and after the Schmitts were taken with the machine scanning at the rate of a line of 70 brackets in 15 seconds. At 5 per word, this would represent 50 wpm. The rate was then speeded up to see if slowness of the response was any problem. Little or no deterioration of the trigger output was noticed until the entire page width was being scanned in something less than 0.5 second. Response time should be no limitation.

Figure 3-3 shows some standard typewritten text as it is picked up by the Lexiphone's photocell array. This sample came from a portable machine with standard cloth ribbon. The ramp output from the timer is displayed vertically on an oscilloscope, and the black pulse signal \( P_b \) from the Lexiphone is applied to the z-axis to modulate the beam intensity. Notice the extremely fine resolution that is characteristic of the 54 cell reading head.
Fig. 3-1  Raw signal from a photodiode. This is taken from one cell as the array scans a row of brackets (((((. Signal is about 50 na p-p, with a minimum of approximately 10 na.

Fig. 3-2  Effect of the Schmitt triggers. Output of the trigger receiving the signal Fig. 3-1.
Fig. 3-3  (a) Magnified view of standard typewriter print.  
(b) Typewritten text as seen by the Lexiphone.
The fine resolution is typical of the 54 cell scanner.
3.1.1 Upper Outline

The top edge decision circuit must produce a pulse when a legitimate white-to-black transition has occurred. Figures 3-4, 3-5, and 3-6 are waveforms taken with the \( P_t \) circuit set to produce a \( P_t \) when a 00011 sequence has occurred. When the page is in the dark, all the cells are indicating black (Figure 3-4) and a top edge pulse is produced at the start of the black island. This is the simplest, most straightforward case, and virtually any criterion would work here. Figure 3-5, top trace, shows a \( P_b \) sequence much more typical of the problems encountered when actually reading a page. The single cell black pulse second from the left could easily be caused by a speck of dust, or dirt from an erasure. In addition, we notice a solitary white cell in the right hand island of black. This could have been due to print imperfections. The \( P_t \) output is shown below. The solitary cells have been ignored, and the correct two top edge pulses have been produced. (\( P_t \) is not produced until the 00011 sequence is encountered and is hence delayed two pulses from the edge it indicates.)

Regardless of the criterion used to determine an upper outline, there will still be certain unavoidable "errors". When passing off the edge of a vertical line, a momentary burst of randomly spaced top edges occurs. The vertical "line" is actually quite irregular when examined closely, and even if it were perfect, errors would still occur when the cells are half over the edge. In this situation they are presented with grey, and due to inherent individual differences, the Schmitts turn off in a random order. Fortunately, this is of a very short duration, and is not significant
Fig. 3-4  **Top edge recognition.** This picture shows the simplest case. There is a huge island of black and only one edge is present. Note that the $P_t$ is not produced until two black cells have been confirmed.

Fig. 3-5  **Spurious pulses** and dropouts are ignored by the $P_t$ circuit. The isolated black cell and the isolated white cell seen in the top trace are ignored and the correct two top edge pulses are produced.
Fig. 3-6 Unavoidable "errors" must still occur. The cells are caught in a random state in the upper waveform. This leads to a random burst of top edges. (this condition exists when the array is passing off the side of a vertical riser)
in the output.

To give an overall check of the machine performance from page to f-function, a set of tests with one to four edges in random arrangements was typed (Figure 3-7). The theoretical value of the f-function is easily determined from these straight line patterns and can be compared to the values calculated by the machine. The results are very good, and verify that the machine is calculating the correct f-function. In the discussion on \( P_t \), an f-function for the "I" was shown. The "a" of Figure 3-11 is a further illustration of the type of f-function actually obtained.

### 3.2 Code Tests

The problem of finding the best aural output for a reading machine has been given much consideration. It is generally agreed that the very nature of a direct translation machine introduces limitations on the attainable reading rates. To give a general idea of how the Lexiphone code compares with other codes, some simple tests were run. The testing pattern used by Haskins labs\(^4\) and Nye\(^6\) to test comparative code performances was used. A set of eight simple four letter words are expressed in the code to be tested and presented in a code-response-code-voice format.

The standard of performance is a speech-like code originated at Haskins Labs and nicknamed "Wuhzi".\(^4\) Wuhzi is produced by a human speaker. The formation of Wuhzi is shown in Table 3-1. The learning curve for Wuhzi forms a basis for comparison. It is interesting to note that the Wuhzi learning curve obtained here is identical to that of Nye.\(^6\) This gives us at least some confidence that the conditions of the test were duplicated.
Fig. 3-7 Tests of the f-computing circuitry. (a) as typed, (b) recorded through the reading head. These straight line patterns of one, two, three, and four top edges test the generation of the \( f \)-function from the page to the end of the \( f \)-computer. The right answers are easy to calculate and can be quickly compared to the measured ones. (see the following pictures)
Fig. 3-8 Measured f-function from test #1

Fig. 3-9 Measured f-function from test #2
Fig. 3-10 Measured f-function, test #3. These three tests provide a final test of the operation of the bulk of the machine. The proper production of these f-functions not only requires that the f-computer be functioning correctly but also those parts of the machine supplying inputs to the f-computer (illumination, scanning, focus, array, Schmitts).

Fig. 3-11 Sample of practical f-function. This f-function of "a" is typical of those obtained during use of the machine on print.
The source for the Lexiphone tests was the machine itself. It is felt that this point cannot be overemphasized, as the output of previous machines had been so unpredictable that it was necessary to record a master tape of sounds and use it as the source. In this case, the words of the test were simply typed on a sheet of paper and read by the Lexiphone directly on to the final test tape.

The Morse test was made using international Morse at 50 wpm, the same speed at which the Lexiphone was run. The list was made by sending on an electronic key at 25 wpm, and running the tape at double speed. This was necessary because 50 wpm is a very high speed indeed to send Morse.

The results of the tests were more or less as expected (Figure 3-12). The Wuhzi test was in surprisingly close agreement with the curve of Nye et al. The Lexiphone is, not surprisingly, somewhat below Wuhzi. The important point is that the Morse is markedly lower than the Lexiphone at this speed of 50 wpm. This bears out the proposal put forth earlier in the thesis that the Lexiphone code (at least at higher speeds) certainly does not have the same number of psychologically perceived elements as Morse, and that decoding can, and does, proceed at a speed much above the Buzz Rate suggested by Cooper and Zahl. It is interesting to note the result when two of the blind subjects already using the machine were asked to try the Lexiphone test. One missed a word on the first try, but after that the scores were perfect. The fact that the correct answers could be called off without hesitation by these expert girls shows, at least, that sufficient information is present for the learning curve to reach perfect score. In fact, the Lexiphone curve
Table 3-1  Transliterations used to produce Wuhzi

<table>
<thead>
<tr>
<th>word</th>
<th>wuhzi</th>
<th>pronounced</th>
</tr>
</thead>
<tbody>
<tr>
<td>with</td>
<td>yekw</td>
<td>yeekwuh</td>
</tr>
<tr>
<td>will</td>
<td>yemm</td>
<td>yeem</td>
</tr>
<tr>
<td>were</td>
<td>yini</td>
<td>yini (tip)</td>
</tr>
<tr>
<td>from</td>
<td>snal</td>
<td>snal (tap)</td>
</tr>
<tr>
<td>been</td>
<td>jiir</td>
<td>jeer</td>
</tr>
<tr>
<td>have</td>
<td>wozi</td>
<td>wuhzi</td>
</tr>
<tr>
<td>this</td>
<td>kweef</td>
<td>kweef</td>
</tr>
<tr>
<td>that</td>
<td>kwok</td>
<td>kwuck</td>
</tr>
</tbody>
</table>

Table 3-2  Wuhzi equivalents of the words used in the 8-word test
Fig. 3-12 Comparison of Code
was carried beyond that shown on the graph and several perfect scores were recorded among group members, although the average was only up to 20 after 18 trials.

Two important points emerge from these code tests. First, by comparison with the Morse, we suggest that the Lexiphone code has fewer psychologically perceived elements than had been assumed by other workers, and is hence amenable to higher speed decoding. Second, and more important from the point of view of this thesis, a learning curve roughly the same as the best codes tested at Haskins Laboratories (Optophone curve) has been obtained with a set of tests recorded directly from the machine. The consistency of the code sounds from presentation to presentation is here due to the dependable performance of the machine itself, and not to the certain repeatability of hearing the same tape recorded sounds over and over.

3.3 Reading Results with the Machine

During the design and construction of the machine, a group of subjects was trained using a computer simulation of the Lexiphone code. The simulation was quite approximate in that it was made with data from relatively coarse manual measurement of the letters. The simulation had a quantized sound, and, as was discovered when the subjects were exposed to the real time machine, certain errors had been made in the original letter measurements. Nevertheless, the sounds were basically the Lexiphone, and it is felt that their experience with the simulation prepared the subjects well for the task of learning to read with the hardware machine.

At the time of writing, only two subjects have had a significant exposure to the machine. They are making very encouraging
progress (Figure 3-13) in reading speed. (c.f. Figure 3-14) By way of comparison, an extensive, 200-lesson training program was developed by the Battelle Memorial Institute for instruction in the use of the Battelle Optophone. At the end of fiscal year 1963, when the final report was given, they say of their top student in the class of '63: "On the test of comparable first grade level material at the conclusion of training, subject A read 21.9 wpm including line change time, and 25.4 wpm excluding line change time". The present Lexiphone is reading grade III material at 30 wpm and the rates are still improving.

3.4 Continuing Work

The best method of assessing the present design and of suggesting any refinements is testing by blind users. To this end, work is under way on 10 Lexiphone machines, which are to be given out for evaluation. At the time of writing, the electronic sections of these machines are nearing completion. The new machines have a much higher parts density, and can be fitted, exclusive of the power supply into less than a 6" cube. Such compact units require great care in the layout and wiring.

Still to be constructed are suitable optical and mechanical systems for the operational machines. Such features as the ability to track in a straight line and at a constant speed, and the ability to change scale, are extremely desirable. Such a system must also be capable of providing strong illumination on the page. In spite of these many necessary features, the reading head must be small. The problems of tracking and variable scale have been handled successfully by earlier workers and it is possible to adopt proven techniques in
Fig. 3-13 Reading Progress with the Lexiphone
Fig. 3.14 Best Two Subjects of Battelle FY 1963 Evaluation Class. Each test represents 10 hours of instruction.
these areas. The Lexiphone, however, has particularly demanding illumination requirements, which may be hard to meet while still maintaining a small physical size.

One possible solution would be to illuminate the page via fiber optics. This would mean the high-intensity light could be located in the fixed part of the machine. Losses, primarily in getting light into a light guide, would increase the intensity of the lamp required, but since it would no longer be part of the reading head, this may not be a serious problem.

In considering the fiber optics, the possibility of transmitting the image itself through a coherent bundle arises. This would make possible a further reduction in the size of the reading head itself. Requirements have been given to the fiber optics division of Bausch and Lomb and they are preparing a device with a coherent inner bundle for image transmission and a non-coherent outer bundle for illumination.

It is suggested that page scanning be done with a handheld probe guided by a mechanical tracking aid. This arrangement is smaller and more versatile than the present moving platform page scanner. In addition, it enables the user to slow down easily over unfamiliar words, and to quicken the pace over familiar text. With such easy control, the beginning user can develop his own techniques of distinguishing similar letters or decoding difficult words.

When the electronics, page scanner, and optics are ready, people must be trained in the use of these machines. At present, work is proceeding, in collaboration with the Faculty of Education, on development of a suitable training program for users of the Lexiphone. Prepared by such a course, they will be in a position to
evaluate the merits and failings of the machine.

To summarise, the following items must be taken care of to obtain the final results of the Lexiphone project:

The present circuit must be packaged into a compact unit. The first of ten such units is now almost completed.

A suitable optical system must be built. It must be compatible with a reading head of small physical size, and provide sufficient page illumination. Provision for scale change (variable magnification) is desirable. The possibility of using fiber optics is being investigated.

A thorough, professionally managed, training program is essential to meaningful results. Such a program is now being developed to train users of the 10 test machines.

Finally, the 10 machines may be distributed to their users, and the training-evaluation program begun. The results of this will guide the course of future work.
4. SUMMARY AND CONCLUSIONS

Early reading machines met with such meagre success that many workers concluded that simple reading devices could never offer satisfactory reading rates. This thesis reviews the older machines and suggests that, while there are fundamental limitations, the real failing of previous machines was in lack of repeatability and an over-dependence on alignment. It is felt that critics of the direct translation machine have been discouraged by calculating theoretical maxima in reading speed that were somewhat low, and by being overly demanding in their estimates of an absolute minimum "useful" speed (Cooper\(^{(12)}\) sets the absolute minimum at 50-60 wpm and the value for an average good user at 100 wpm). Beddoes\(^{(1)}\) has proposed a code (Lexiphone code) which is not dependent on alignment, allowing a considerable latitude in tracking, and which provides a constant clue (mean pitch) to the correct alignment. Tests with computer simulations showed that this code could be learned at a rate fast enough to warrant building a machine. This thesis has described the design and construction of a machine to produce the Lexiphone code dependably and repeatably. A high resolution (54 cell) scanner has been used to take full advantage of the position independence inherent in the Lexiphone code. Digital circuits have been used throughout the code generating unit, and a high degree of dependability, as well as a physical compactness, has been attained. The prototype machine has received extensive use over the past three months without requiring adjustment. Reading results with this prototype (Section 3.3) have shown it to be already markedly superior to the Optophone, and performance has been steadily improving.
A speech-like code (Wuhzi), the Lexiphone code, and Morse code were subjected to an eight word screening test used to compare code performances in the past. The Wuhzi curve corresponded exactly with the curves of other workers. The Lexiphone was roughly the same as the Optophone (as tested at Haskins Labs), and the Morse was well below both. By comparison with the Morse, it was suggested that at the speed of these tests (50 wpm), the machine code has less psychologically perceived elements than Morse, and hence lends itself to more rapid assimilation. The close correspondence between the Lexiphone and Optophone learning curves served to confirm the suggestion put forth earlier in the thesis that the true failing of simple machines had not been in the code per se, but rather in the capability to produce the code with a real machine manned by a blind operator. The learning curves obtained with artificially repeated sounds would indicate the Optophone could have been learned at 50 wpm, whereas this value was not approached in practice.

It is concluded that the Lexiphone is well suited to realize the potential speed of the code itself, and that this value is above the somewhat pessimistic values calculated previously.

4.1 Remaining Questions

Net reading speeds of 30 wpm are being obtained on Grade III material with the present Lexiphone. It is felt that it is not unreasonable to expect 40 wpm as a maximum. It is pointed out that such speeds are near the upper limit for unprocessed spelled speech (each letter spoken). Golden showed that, by use of a time compressor, code sounds could be understood at a much faster rate than
the uncompressed version. He also optimized the parameters for this compression. The effect of compression on the Lexiphone code from the machine deserves further consideration. If results similar to Golden’s are obtained, it could make the Lexiphone-compressor unit comparable in reading speed to the obvious next step in complexity—the character recogniser with a spelled-speech output. A prototype (Cognodictor) of such a machine is being tried by Mauch Laboratories (18). However, the tracking tolerances are very small and reading rates only of the order of 25 wpm have been recorded. The advantage of the recognition machine is, of course, that it requires very little practice to use, but if similar reading rates can be obtained with the much smaller, less critical machine, it would warrant spending time learning the code. More research on code compression could make this possible.

We also should ask if the present Lexiphone code is the best. Reading performance has shown the Lexiphone to be superior to the Optophone (Section 3.3) but it is not known precisely why. There is room here for more investigation of the codes. Theoretical work on the two codes, and theoretical predictions of reading rates, may reveal a better code to use. Work towards a model of code perception would be very helpful in determining the optimum codes for future machines.

It is desirable that the reading machine be portable and of "reasonable" price. The present Lexiphone meets the small size requirement, but component cost is approximately $600. While the primary consideration is a machine that works well, it is important to determine if the present machine could be produced at a "reasonable" price.
APPENDIX

Logic Level Symbology

- Ground Level for Assertion
- Positive Level for Assertion
- Pulse to Positive Level for Assertion
- Pulse to Ground Level for Assertion
- Nonstandard Signal
REFERENCES


