PROCEEDINGS

1973 CARNAHAN CONFERENCE ON ELECTRONIC PROSTHETICS

UKY BU104
NOVEMBER 1973

Conference Sponsors:

University of Kentucky
College of Engineering
Department of Electrical Engineering
Office of Engineering Continuing Education
Lexington Section IEEE

In Cooperation With:

Technical Groups of the Institute of Electrical and Electronics Engineers:
- Computer Society
- Control System Society
- Engineering in Medicine and Biology Technical Group

ENGINEERING UK

UNIVERSITY OF KENTUCKY 1865 UNITED WE STAND DIVIDED WE FALL

For additional copies or information address correspondence to:
ORES Publications
College of Engineering
University of Kentucky
Lexington, Kentucky 40506

A PUBLICATION OF THE OFFICE OF RESEARCH AND ENGINEERING SERVICES

UKY BUL 104

PROCEEDINGS

1973 CARNAHAN CONFERENCE ON ELECTRONIC PROSTHETICS

NOVEMBER 1973

John S. Jackson, Editor

R. William DeVore, Publishing Editor

Conference Sponsors:

University of Kentucky
College of Engineering
Department of Electrical Engineering
Office of Engineering Continuing Education
Lexington Section IEEE

In Cooperation With:

Technical Groups of the Institute of Electrical and Electronics Engineers:
- Computer Society
- Control System Society
- Engineering in Medicine and Biology Technical Group

DEDICATION

This record of the PROCEEDINGS of the 1973 CARNAHAN CONFERENCE on ELECTRONIC PROSTHETICS is dedicated to two men whose professional accomplishments have so greatly benefitted the handicapped.

James F. Garrett
Assistant Administrator
Social and Rehabilitation Services
HEW
Washington, D.C.

Capt. G.E. Welch
Director, Access for the Disabled
Central Council for the Disabled
London, England

CONTENTS

Page

WEDNESDAY, SEPTEMBER 19, 1973
FIRST SESSION: DEAF ORAL PROSTHESES

CUTANEOUS STIMULATION FOR SOUND LOCALIZATION AIDS
J. L. Mason, and
N.A.M. Mackay
Queen's University
Kingston, Ontario, Canada 1

THE SAMPLING OF SOUNDS AS A MEANS OF MAKING SPEECH INTELLIGIBLE TO THE DEAF: THEORIES AND RESULTS
Emanuele Biondi
Istituto di Elettrotecnica de Elettronica
del Politecnico di Milano
Centro Teoria dei Sistemi, C.N.R.
Italy 5

SECOND SESSION: DEAF-ORAL PROSTHESES

THE VSTA: AN APPROACH TO THE SPEECH TRAINING INSTRUMENTATION PROBLEM
L. C. Stewart,
Robert A. Houde, and
Wilbur D. Larkin
Center for Communications Research
50 West Main Street
Rochester, New York 14614 10

A COMPUTER BASED VISUAL AID FOR THE DEAF
Stanley C. Person, and
Ian B. Thomas
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts 01002 15

A MULTIPARAMETER VISUAL SPEECH DISPLAY
Vassilios J. Georgiou, and
Ian B. Thomas
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts 01002 24

BANQUET ADDRESS
Captain G.E. Welch
Director, Access for the Disabled
Central Council for the Disabled
London, England 28

THURSDAY, SEPTEMBER 20, 1973
THIRD SESSION: CEREBRAL PALSY REHABILITATION DEVICES AND TECHNIQUES

APPLICATION OF ELECTRONIC PROSTHESES IN REHABILITATION AND EDUCATION OF MULTIPLY HANDICAPPED CHILDREN
F.A. Harris,
F.A. Spelman,
J.W. Hymer, and
C.E. Chase
Department of Physiology and Biophysics
Regional Primate Research Center
University of Washington Medical School
Seattle, Washington 33

CONTENTS (continued)

Page

TYPEWRITER FOR TEACHING SEVERELY HANDICAPPED CHILDREN
P. J. R. Harding
A. J. Kozak
K and H Engineering, Ltd.
Edmonton, Alberta, Canada
Y. J. Kingma
K. A. Stromsmoe
J. F. Vaneldik
Electrical Engineering Department
University of Alberta
Edmonton, Alberta, Canada 43

THE AUTO-MONITORING TECHNIQUE AND ITS APPLICATION IN THE AUTO-MONITORING COMMUNICATION BOARD (AUTO-COM), A NEW COMMUNICATION DEVICE FOR THE SEVERELY HANDICAPPED
Greg C. Vanderheiden,
Andrew M. Volk, and
C. Daniel Geisler
Cerebral Palsy Communication Group
Electrical Engineering Department
University of Wisconsin
Madison, Wisconsin 53706 47

FOURTH SESSION: READING AND MOBILITY AIDS FOR THE BLIND

THE TRANSICON
Hillel Bar Lev
Consultant
The Scientific Research Foundation
Jerusalem 52

SEARCH AND RETRIEVAL TECHNIQUES FOR USE WITH TAPE RECORDINGS
Emerson Foulke
Perceptual Alternatives Laboratory
University of Louisville
Louisville, Kentucky 40208 67

THE STEREOTONER READING AID FOR THE BLIND, A PROGRESS REPORT
Glendon C. Smith
Mauch Laboratories, Inc.
3035 Dryden Road
Dayton, Ohio 45439 74

FRIDAY, SEPTEMBER 21, 1973
FIFTH SESSION: AIDS FOR MUSCULAR AND NEURAL IMPAIRMENT

THE NEW C-5 LASAR CANE FOR THE BLIND
J. Malvern Benjamin, Jr.
Bionic Instruments, Inc.
221 Rock Hill Road
Bala Cynwyd, Pennsylvania 19004 77

STUDIES IN GAIT IMPROVEMENT BY FUNCTIONAL STIMULATION OF THE LATERAL POPLITIAL NERVE
J. Vredenbregt
Institute for Perception Research
Eindhoven, Netherlands 83

CONTENTS (continued)

Page

A MINIMUM ENERGY EXOSKELETON
Eugene V. Kosso
Electrical Engineering Department
University of Nevada
Reno, Nevada 89507 86

STOCHASTIC ANALYSIS OF EMG SIGNALS FOR MULTI-FUNCTIONAL PROSTHESIS CONTROL PURPOSES
Daniel Graupe,
William K. Cline, and
Ted K. Kaplon
Department of Electrical Engineering
Colorado State University
Fort Collins, Colorado 80521 90

DEVELOPMENT OF A PROGRAMMABLE, RESPONSIVE CARDIAC PACING SYSTEM
Reese S. Terry, Jr.
Cordis Corporation
Miami, Florida 33137 97

SIXTH SESSION: QUADRIPLEGIC AIDS

COMPUTER IMPLEMENTATION OF A PARAMETRIC MODEL FOR BIPED LOCOMOTION KINEMATICS
Thomas Charles Hartrum
The Ohio State University, and
U.S. Army Electronics Command
Eatontown, New Jersey 07724 99

VOCAL SWITCHING AID FOR QUADRIPLEGICS
Robert A. Curtis
Woodridge, Virginia 22191. 119

MOBILITY AID FOR QUADRIPLEGICS
Joachim A. Maass, and
Vera S. Maass
University of Missouri
Kansas City, Missouri 64128 123

MOBILITY OF THE BLIND*
Howard Freiberger
Veterans Administration
252 Seventh Avenue
New York, New York 10001 126

A THIRD GENERATION SELF-TIMING PAGE MARKER*
Deane B. Blazie
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, Maryland 21005 130

* From Fourth Session
Submitted after publishing deadline

CUTANEOUS STIMULATION FOR SOUND LOCALIZATION AIDS

by

J.L. Mason and N.A.M. Mackay
Queen's University
Kingston, Ont., Canada

Summary. The paper discusses the results of experiments carried out to determine the feasibility of using electrocutaneous stimulation in sensory substitution systems in general and sound localization aids in particular.

Introduction

Recent advances in electronic technology have made possible the development of sensory aids for the deaf that provide some facility for sound localization and a primitive form of speech perception. These devices sample the auditory field and present to the user some function of the field in the form of visual, vibrational or electro-cutaneous patterns. Because of the relatively simple function that sound localizers must perform and the immediate advantage to which they can be used, there has been greater success in the development of these devices than for those which attempt to aid speech perception. The more common sound localizers present their information to the user through skin stimulation: either vibro-tactile or electro-cutaneous. Because of their inherently lower power requirements, smaller size and ease of fitting, electro-cutaneous stimulators would be favoured for portable aids,. but too little is known about the control of the interaction between electrical stimulation and the central nervous system for these devices to be well accepted at present. A study of these effects is being conducted at Queen's University. The results of experiments related to the use of electro-cutaneous stimulation in sensory aids are discussed in this paper.

By carefully choosing the parameters of the electric impulses to the skin, namely, current pulse width, amplitude, duration and repetition rate, it is possible to elicit in the user a sensation of vibration at the point of stimulation. Because the phenomenon was not well understood, however, control of the sensation was unreliable. For example, the parameters required to produce the sensation at different times and in different parts of the body, varied drastically. Often, the feeling of vibration changed to a pricking or unpleasant burning pain with small changes in stimulus parameters. Only if certain types of electrode paste were used, or the skin suitably prepared beforehand, could reliable stimulation be achieved. The use of electro-cutaneous stimulation as the output function of a sensory aid could, therefore, reduce the dynamic range of the device in an uncontrolled manner by causing uncomfortable sensations of pain to the user. A number of experiments have been conducted to isolate these problems. Methods of removing them have been developed and are discussed in the following sections.

A phenomenon that may be of particular use in localization aids is that of the so-called "Phantom sensation" described by Von Bekesy (1), in which two spatially distinct vibro-tactile stimulations may be perceived as a single, fused sensation, the location of which is controlled by the relative intensities of the two original stimuli. This phenomenon has been found to be useful in a sound localization aid in which the intensity of sensation at one of the points of stimulation is related to the sound intensity at a microphone located at one ear of the user (2). Differences in sound intensity at the two ears, due to a sound source located at an angle to the median plane of the head, cause the fused sensation to move, thereby providing a means of localizing the sound. The appearance of a fused sensation when vibro-tactile stimulation is used has been explained by Bekesy (3) on the basis of neural inhibition. Since a similar inhibition can be anticipated when electro-cutaneous stimulation is used, it can be expected that the phenomenon will occur in either modality. A number of experiments to verify this conclusion are also described below.

Experiments on Electro-cutaneous Stimulation

The experiments performed in this study have overlapping features but can be loosely classified into three categories:

Dynamic Range

The dynamic range for this study can be defined as the range in intensity of stimulation required to change the sensation the subject discerns from one of minute vibrations to one of an uncomfortable pain. Although wide variations in dynamic range between subjects were noted, a systematic method of increasing the dynamic range of all subjects was found. Factors of importance were initial skin conditions, the rate of increase of stimulation, and the maximum intensity of stimulation. The effects of electrode paste on the dynamic range were found to be readily predictable.

The Electrical Parameters of the Skin

The electrical parameters of the skin vary greatly with time, changes in skin preparation and stimulation intensity. Changes in the resistance of the corneal layer were found to vary with stimulus current and provided a good measure of impending skin breakdown. Various levels of pain can be associated with physiological changes in the cutaneous layer.

The Phantom Sensation

Experiments indicate that the phantom sensation can be elicited with electrocutaneous stimulation. The parameters of spatial and temporal resolution and the relative intensity of stimuli required at each stimulus position have been measured. These appear to be different from those noted for vibro-tactile stimulation.

Equipment

The equipment shown schematically in Figure 1 was used in the first two series of experiments to apply current pulses, of positive and negative polarity, to the skin. Duplication of some portions of the set-up was required to elicit the phantom sensation.

Bipolar pulses produced by the pulse generator were amplified by means of various voltage and power amplifiers. The magnitude of the forward and reverse current was controlled independently by the current limiter. The limiter output, a series of positive and negative going constant-current pulses, was then directed to coaxial electrodes and applied to the skin. For most of the experiments described the pulse parameters were:

Pulse width: 1mSec Pulse height: .5 to 10mA
Repetition rate: 13Hz
Pos/Neg Pulse Separation: 2 mSec

Silver plated coaxial electrodes were used; a photograph and schematic diagram showing their construction is shown in Figure (2). The panic button allowed the subject to shut off all stimulation instantaneously if necessary. Happily, this feature was never utilized.

The surface resistance of the skin was measured by monitoring the pulse voltage and current at the electrodes. During each pulse, the information was sampled in the sample and hold units, and displayed on a chart recorder, digital volt meters or oscilliscope. A continuous record of skin resistance could therefore be obtained.

Results

Dynamic Range

To achieve some degree of reliability in these measurements, a controlled psychophysical experiment was devised. The subject was briefed on the operating procedure, the electrode strapped to his arm and the stimulation current increased at the rate of 0.8 mA per minute. The subject was asked to indicate with one word when he first felt a sensation. The stimulation was increased in intensity at the constant rate until the subject indicated that he did not wish to have the stimulation increased further; the stimulation current was then held at this level for thirty (30) seconds. The current was then reduced at a constant rate until the subject indicated that sensation had disappeared. The experiment was repeated five times with rest periods of two minutes between each run.

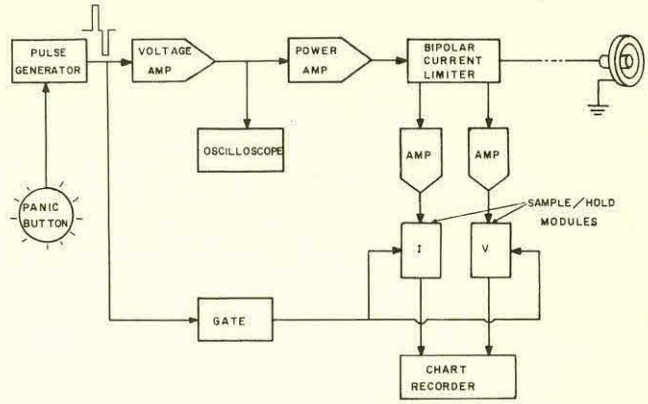


FIGURE 1: BLOCK DIAGRAM OF STIMULATOR.

Twelve subjects were tested; most were tested more than once. The effect of skin preparation was determined by performing tests with and without electrode paste and with the skin prepared with various alcohol and water solutions. The results are summarized below.

(i) The subject's dynamic range of sensation increased in the first few runs then remained constant in further runs. That is, a subject's tolerance to pain and hence the range of comfortable stimulation could be increased providing certain initial conditions described below were met.

In addition, the sensation and pain thresholds for a given point on the arm were consistent over long periods of time. For example, if the point of stable dynamic range was reached, any level of stimulation below the pain threshold could be instantaneously applied to the subject without eliciting a pain sensation, even after rest periods of ten minutes.

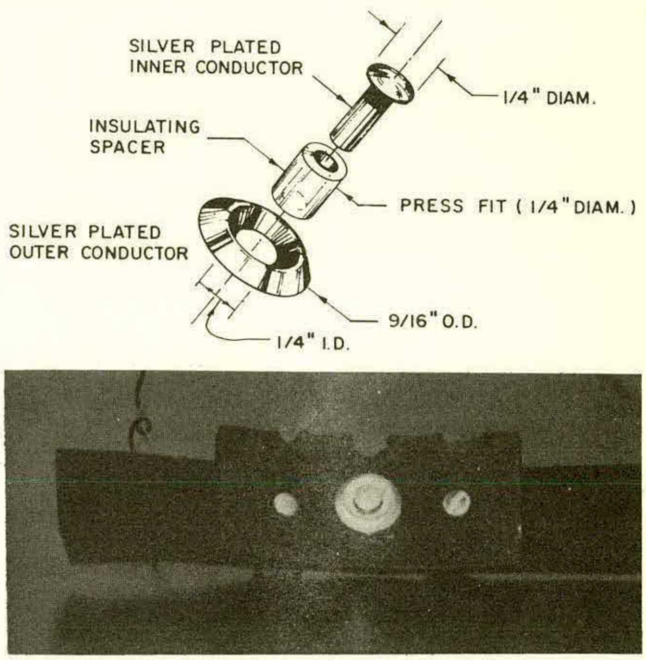


FIGURE 2: THE CONCENTRIC ELECTRODE.

The dynamic range for a particular point on a subject's body was also consistent from day to day. Thus a subject's pain and sensation thresholds for any point on his body can be predicted from initial measurements.

(ii) However, as also noted by Pfeiffer (4) there was a wide variance in sensation and pain thresholds between subjects and between various parts of the forearm for the same subject. Sensation thresholds varied from 0.3 mA to 1.6 mA, pain thresholds from 2.4 mA to 11.2 mA. There also appeared to be no relationship between a subject's pain and sensation thresholds. That is, a knowledge of a subject's pain threshold would not aid in the prediction of his threshold of sensation.

Electrical Parameters of the Skin

Some of the electrical parameters of the skin were closely correlated to the sensation perceived by the subject. The initial condition of the corneum was of major importance in obtaining a comfortable sensation with electro-cutaneous stimulation. This can be demonstrated by Figures 3 and 4, showing plots of measured skin resistance versus time under different operating conditions. In Figure 3, a comparison is made between the resistances of dry skin, skin saturated with electrode paste and skin moist with water. For all three runs the stimulation current was held constant at 1.0 mA. The first portion of the dry-skin curve shows an initial, rapid decrease in resistance followed by a more gradual decrease. The paste curve shows an initial increase in resistance, then again a gradual decrease. The resistance of moist skin begins at a lower level than that of dry skin but approaches the same final value.

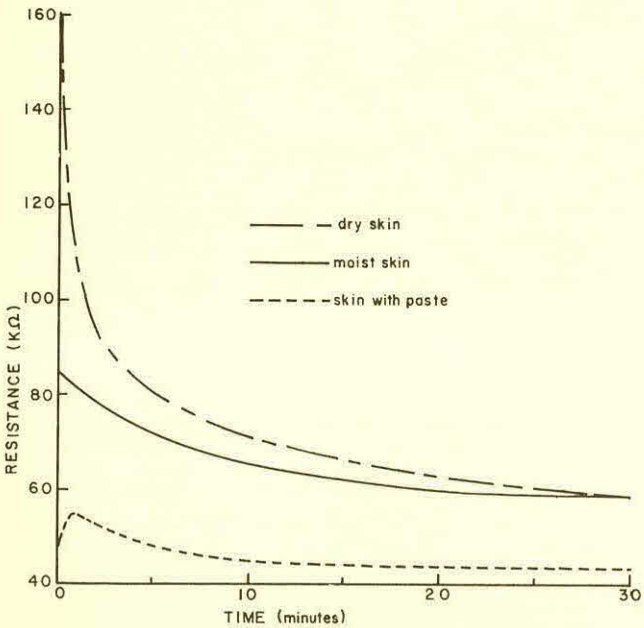


FIGURE 3: EFFECTS OF PREPARATION ON SKIN RESISTANCE.

The gradual changes in resistance can be traced to variations in moisture level in the corneum with time: the inter-electrode resistance increases as the paste and water coated skin dry out but decreases as perspiration develops beneath the electrodes. On the other hand, the initial rapid decrease in resistance of the dry skin curve is related to dielectric breakdown in the corneum. Because of the high corneum resistance, the inter-electrode voltages required to drive even small currents through the skin are sufficiently high to cause dielectric breakdown of the corneum, producing microscopically-small punctures in the outer cutaneous layers. An uncomfortable pricking sensation accompanies the dielectric breakdown but disappears when the resistance assumes a more gradual change with time. To avoid this pricking sensation, therefore, one merely has to ensure that the skin is moist before attaching the electrodes. Once the electrodes are in place, perspiration will keep the inter-electrode resistance at a low enough level to avoid breakdown. If this procedure is followed, electrode pastes will not be required.

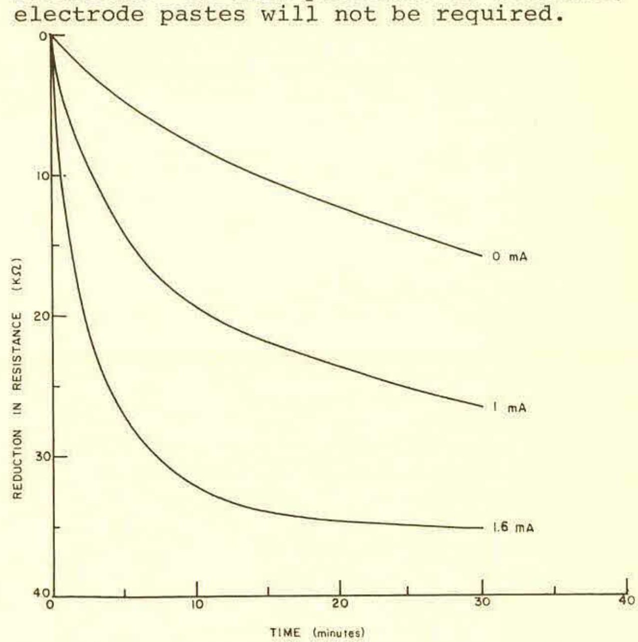


FIGURE 4: EFFECTS OF STIMULUS INTENSITY ON SKIN RESISTANCE.

Figure 4 shows the relationship between inter-electrode resistance and intensity of stimulation. The decrease in resistance both with applied current and time appears to be related to changes in the physiological parameters of the skin as evidenced by an area of erythemia beneath the electrodes which increases with stimulation. These changes influence the dynamic range of sensation; that is, a subject's pain threshold can be increased to some degree if the level of stimulation is increased slowly enough to allow these changes to take place. This has been borne out in the results of the tests mentioned earlier in which a subject's dynamic range would increase in the first few test runs but stabilize with time.

The Phantom Sensation

A single fused sensation of vibration can indeed be produced on the forearm by properly controlling the parameters of electrical stimulus at two loci. The phenomenon has similar features to those observed when vibro-tactile stimulation is used. The apparent fused sensation can readily be made to move between the electrodes by varying the ratio of current applied at each electrode but reproducible results can only be obtained after the electrodes have been attached for a 'settling-in' period. In addition, resolution appears to be lower for electric stimulation but is dependent on electrode spacing and varies from subject to subject. Typically, subjects can resolve an apparent sensation to within half an inch when electrodes are spaced eight inches apart. Eight inches is the average maximum spacing of electrodes at which a fused sensation can be elicited.

Conclusions

The foregoing experiments indicate that, although there are still many unanswered factors, electro-cutaneous stimulation can be used to advantage in sensory substitution systems and in particular, in sound localization aids. By following a well-defined procedure, it is possible to produce a comfortable vibrational sensation with no apparent short-term injury to the body. The possibility of using only two electrodes to vary the position of the apparent sensation offers a further hopeful sign to the future of electro-cutaneous stimulation in sensory aids.

References

1. Bekesy, G. Von, Psych. Rev., 66, 1, 1959.

2. Saunders, F.A., I.E.E.E. Trans. Aud. & Electroacoustics, Au-21, 285, 1973.

3. Bekesy, G. Von, J. Opt. Soc. Am., 50, 1060, 1960.

4. Pfeiffer, E.A., Med. & Biol. Engng. 6 637, 1968.

THE SAMPLING OF SOUNDS AS A MEANS OF MAKING SPEECH INTELLIGIBLE TO THE DEAF: THEORIES AND RESULTS*

by

Emanuele Biondi
Istituto di Elettrotecnica ed Elettronica
del Politecnico di Milano
Centro Teoria dei Sistemi, C.N.R.

Leonardo Biondi

Summary. A new method to make human speech intelligible for the profoundly deaf is presented. This method consists primarily on the delivery of sampled sound. Its effectiveness is explained with two different theories, therefore some results of tests are given.

FOREWORD

Many deaf persons, including those suffering from severe impairment, have some residual hearing at low frequencies (for instance, between 200 and 1000 Hz) and exhibit more serious loss at higher frequencies. Since the frequency band which retains a remainder of sensitivity is very narrow in these cases, it is impossible, or at any rate extremely difficult, to make speech intelligible by means of ordinary amplifying aids. In fact, it is well known that speech consists of phonemes which contain significant frequencies beyond such sensitivity band.

Over a number of years several methods have been suggested in order to overcome this difficulty and appliances have been developed, many of which operate basically on the principle of frequency transposition (for single or multiple bands).

A critical analysis of the criteria underlying all developments to date and a study of a mathematical model, albeit approximated and qualitative, of the auditory system suggested an entirely different process. In this paper some research initiated many years ago are reported (1). These lead to the realization of a new system making speech intelligible to the profoundly deaf. This method is essentially based on the delivery of sampled sound to the patient.

In the next section, some general considerations that guided the authors in the preliminary studies are reported, and in addition, the system and the corresponding supporting theories (transposition theory and pseudo-carrier wave theory) are described. As far as the second theory is concerned, the present exposition is substantially different than that previously presented but has the same basic concept. Finally, the results of some experiments are reported.

General Considerations (2)

The optimal design problem of direct prostheses for profoundly deaf people is reduced to that of determining how to process the natural input signals in order to maximize the amount of information contained in the output of the prosthesis-impaired sense organ system.

If we consider the scheme in Fig. 1, we can define at least two errors, which can characterize the difference between the behaviour of the normal and impaired sense organ. For the sake of simplicity it is assumed that the lesion is located at the level of the transmission system.

The error E_1 might be determined. In fact, its determination requires the knowledge of models of the peripheral part of the system, for which sufficient experimental data are generally available. Unfortunately, the error E_1 is not the most significant one. Prostheses designed to minimize E_1 can introduce a coding of the external messages such that the central nervous system is then completely unable to carry on the task of decoding and interpreting the messages. The failure of some acoustical prostheses based on the minimization of E_1 may be explained in this way.

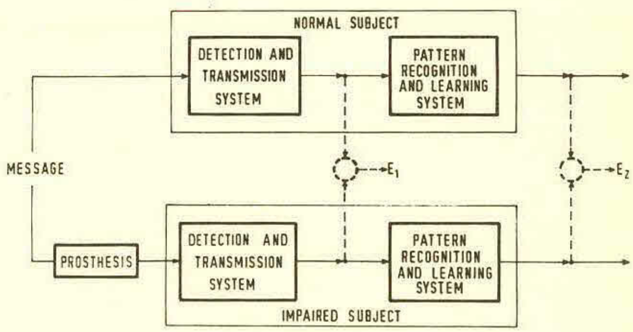


Fig. 1. Scheme for the problem of acoustical prosthesis.

The error E_2 is undoubtedly the most important one, since it defines the difference between the pattern recognition made by the normal and the impaired subject. At present, no satisfactory models are available for estimating E_2. However, on the basis of the present knowledge the following trial and error process can be started.

An incomplete mathematical model of the sense organ is first constructed, then a prosthesis can be designed in order to minimize E_2 according to the model. As long as new pieces of information about the system behavior are obtained, the characteristics of the model and those of the prosthetic device are modified accordingly. The experiments which can be planned by the use

* This research was financed by the "Biondo Biondi-Enrico Romagnoli" Foundation and by Italian National Research Council (C.N.R.)

of the prosthesis may be of great help in obtaining a better knowledge of the system.

Also, the experiments made by using direct prostheses on normal subjects may be extremely important. As a matter of fact, any manipulation of the natural messages introduced to facilitate understanding by an impaired subject, makes understanding more difficult for a normal subject. On the other hand, experiments tend to suggest that, in general, if a normal subject is unable to understand the new coded messages after a brief period of training, the impaired subject will fail also.

An approach to the study of unconventional prostheses for impaired sense organs, based on the use of mathematical models, can be summarized in the following steps:

1. Set up a mathematical model of the sense organ considered. Fven if roughly approximated, the model can suggest new types of direct prostheses.

2. Test the new prosthesis on normal subjects. Prostheses not passing these tests should be rejected, since they will probably fail also with impaired subjects. In this case, turn back to point 1 and try to improve the model by going deeply inside the anatomy and physiology of the sense organ considered. Otherwise,

3. Test the prostheses on impaired subjects. Fither positive or negative results can be used to improve the preliminary model of the system. Also, the classification of the impaired subjects can be improved by interpreting the results in terms of structure and parameters of the model.

4. If an improvement of the mathematical model is obtained, try to modify the prosthesis accordingly, and repeat the procedure from point 2.

THE SAMPLING SYSTEM AND ITS POSSIBLE THEORETICAL JUSTIFICATION

According to these considerations, seven years ago we conceived an apparatus for profoundly deaf people, based on the well known sampling technique.

We shall now summarize a few basic ideas concerning the theory of sampled data systems. For the symbology we refer the reader to the diagram in Fig. 2 and to the oscillograms in Fig. 3. Let:

$$e(t) = \sum_{1}^{M}{}_{n} V_n \cos\left(n \frac{2\pi}{T_s} t + \theta_n\right)$$

be a periodic input signal (of period T_s). Assuming that the appliance is provided with an ideal sampler (with a sampling frequency $f_c = 1/T_c$) the signal delivered to the output of the hold circuit will be u(t) as shown in Fig. 3.

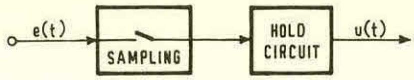


Fig. 2. Scheme of sampled data system.

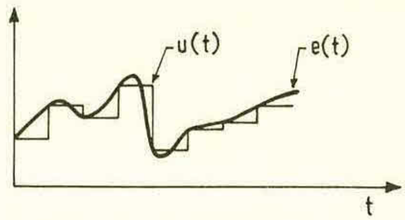


Fig. 3. Signals entering the sampler and coming out of the zero order hold circuit.

If we assume that the hold circuit is an ideal low-pass filter, then:

$$u(t) \simeq \sum_{k=-\infty}^{+\infty} \sum_{n=1}^{M} V_n \cos\left[\left(n \frac{2\pi}{T_s} + k \frac{2\pi}{T_c}\right)\left(t - \frac{T_c}{2}\right) + \theta_n\right] \qquad (1)$$

with the following limitation:

$$\left| \frac{n}{T_s} + \frac{k}{T_c} \right| \leq \frac{1}{T_c} \qquad (2)$$

In effect, owing to the real operation of the hold circuit, u(t) also contains harmonics which eventually will not satisfy (2).

From equation (1) it is easy to recognize the well known fact that multiple frequency transposition of the incoming signal, with preservation of the direct signal, is effected on a single-channel apparatus by sampling.

After the apparatus was developed, the relevant tests were performed, with sampling frequencies lying between 1000 and 4000 Hz, according to the subjects being tested and in part also according to their degree of achieved proficiency.

The following two facts were revealed by these tests:

a) Sampled natural speech is readily intelligible to a normal subject. In other words, distortion of the input signal as a result of sampling is small. This is also explained intuitively by the fact that the oscillogram of u(t) reproduces the incoming signal e(t) at any rate in its microscopic aspects, which probably are alone essential to understanding.

b) After more or less protracted training, with results varying from one subject to another, sampling makes speech intelligible even to the profoundly deaf (see the next point).

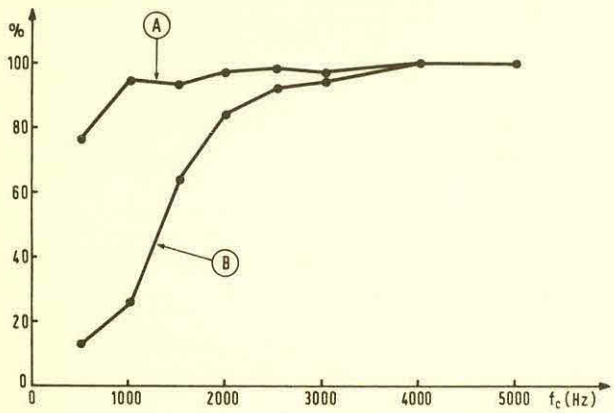


Fig. 4. Scores for intelligibility of sampled words for normal persons (f_c is the sampling frequency).

In Fig. 4 we report the results of the experiments made for normal persons. The experiment consists of recognizing 20 Italian words, quite similar and known to 10 normal subjects, at different sampling frequencies. The subjects were divided in two groups. The scores of curve A are concerned with 5 subjects and the sampling frequency was going down; the scores of curve B are concerned with the other 5 persons and the sampling frequency was rising. It is important to note that the tests themselves were a good training for the subjects of curve A.

In comparison with the other transposition methods, in favor of the sampling equipment, there is the experimental remark according which, by this method, the transpositions are made in such a way that the signal remains intelligible for the normal persons, and this is achieved probably because the transpositions are made without substantial changing of the signal in the time domain (see Fig. 3).

The favorable results may be justified by an alternative theory, the pseudo-carrier wave theory, and it is now presented in a new version.

Let us first review the results of some neurophysiological experiments performed on the auditory nerve of cats (3). The histograms of intervals between spikes may be interpreted as follows:

a) In the absence of stimulus, a statical distribution of spikes of the type as reported in Fig. 5a can be noticed, with the exception on the shortest

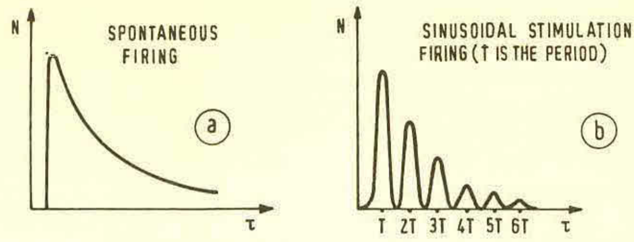


Fig. 5. Interval histogram (τ is the duration of the intervals; $N(\tau)$ is the probability density function).

interspike intervals τ, a Poisson type distribution is shown (4).

b) In the presence of a sinusoidal stimulus (less than 4000 Hz), even at low amplification, the histograms present the pattern reported in Fig. 5b. This phenomenon has been referred to as "synchronization." It is a widespread opinion that synchronization could be the code mechanism by which the central nervous system recognizes frequencies.*

Let us now consider Fig. 6 where $e_N(t)$ is a white noise, the spectrum of which is limited by a filter having a corner frequency f_N; this noise signal at amplification

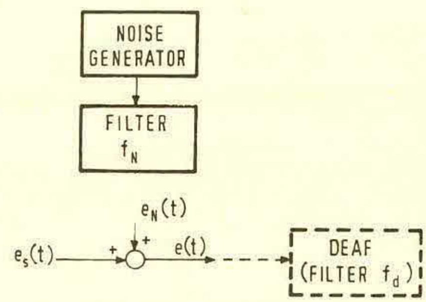


Fig. 6. General pseudo-carrier apparatus.

A_N is sent to the subject together with a sinusoidal signal $e_s(t)$ of frequency f_s and amplification A_s. **

Let f_d be the highest frequency intelligible by the

* The diagram of Fig. 5b is typical for frequency values between 200 and 1000 Hz. For lower frequencies and for frequencies between 1000 and 4000 Hz, the histograms are slightly different but do not affect the base of our considerations. For frequencies higher than 4000 Hz, the mentioned phenomenon vanishes.

** Notice that the described system in Fig. 6 can be considered as the principal scheme of other types of apparatus. The pseudo-portant theory might also explain the good results achieved by the clipped speech apparatus (5) for the deaf with "a high band" filter characteristic.

deaf and we assume that $f_d = f_N$.

If $f_s > f_d$, the signal $e_s(t)$ alone is not perceived by the subject. As far as the signal $e_N(t)$ is concerned, since $f_N = f_d$, there exists a value A_N of the amplification which makes the noise detectable by the deaf.

It is easy to prove experimentally that, if the signal $e(t) = e_N(t) + e_s(t)$ is sent with amplification $A''_N < A'_N$, this signal becomes detectable, moreover, probably the deaf person perceives a sensation like that received by a normal person stimulated with the sine wave alone. Fig. 7 presents a heuristic demonstration of the preceding statement. For the sake of simplicity we assumed that the thresholds for spike generation (due to the mechanical displacement of basilar membrane) for a normal subject are the same in every point of the membrane, whereas the same threshold varies for a deaf person as shown in the figure. For simplicity of exposition, the displacements of the membrane for the signal e(t) are supposed as a sum of the displacements due to the signals $e_N(t)$ and $e_s(t)$.

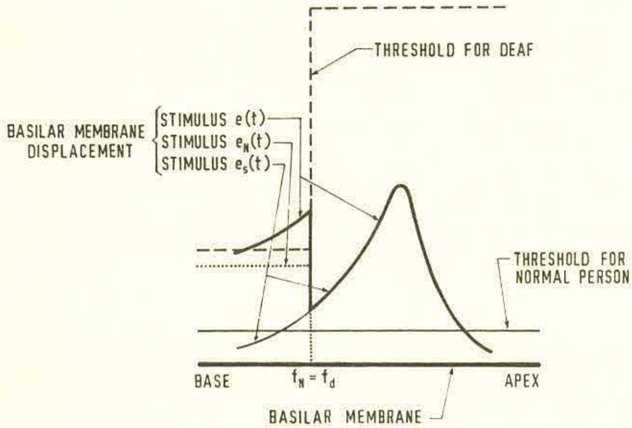


Fig. 7. Schematic demonstration of pseudo-carrier theory.

Let us introduce two hypotheses:

a) the synchronization is a code procedure (may be not the only one) by which the central nervous system recognizes the frequencies,

b) by using suitable values of A''_N and A_s the above mentioned synchronization at the frequency f_s occurs in some fibers, caused by the signal e(t).

From these hypotheses, it follows that a deaf person detects the signal e(t) and may have the same sensation as a normal person when he receives the signal $e_s(t)$. Thus, the noise $e_N(t)$ is used as a carrier to allow the transpositions at neuron level. We could have the same results also if the signal $e_s(t)$ is human speech.

It is now easy to correlate the sampling system to the above mentioned facts. It can be easily proved theoretically and experimentally that the sampling of an acoustical signal generates remarkable low frequency noise.

As an example, in Fig. 8 the power spectrum of the Italian "i" obtained without sampling and with the sampling at 1000 Hz are reported.

In conclusion one can argue that the sampling method exploits both the transposition effect and pseudo-carrier wave effect (or transposition at neuron's level).

USE OF THE EQUIPMENT AND RESULTS WHICH MAY BE OBTAINED.

The experiments conducted up to now, not only directly by us but by private individuals and in schools for the deaf, have given results which in general may be defined satisfactory, at least for an important portion of the subjects under examination. We are going to explain qualitatively the results of our direct tests. For some other results, quantitative ones, see (6). We are now preparing a new report with the last results.

The subjects examined can be subdivided, according to the use of the sampled signal prosthesis, into two

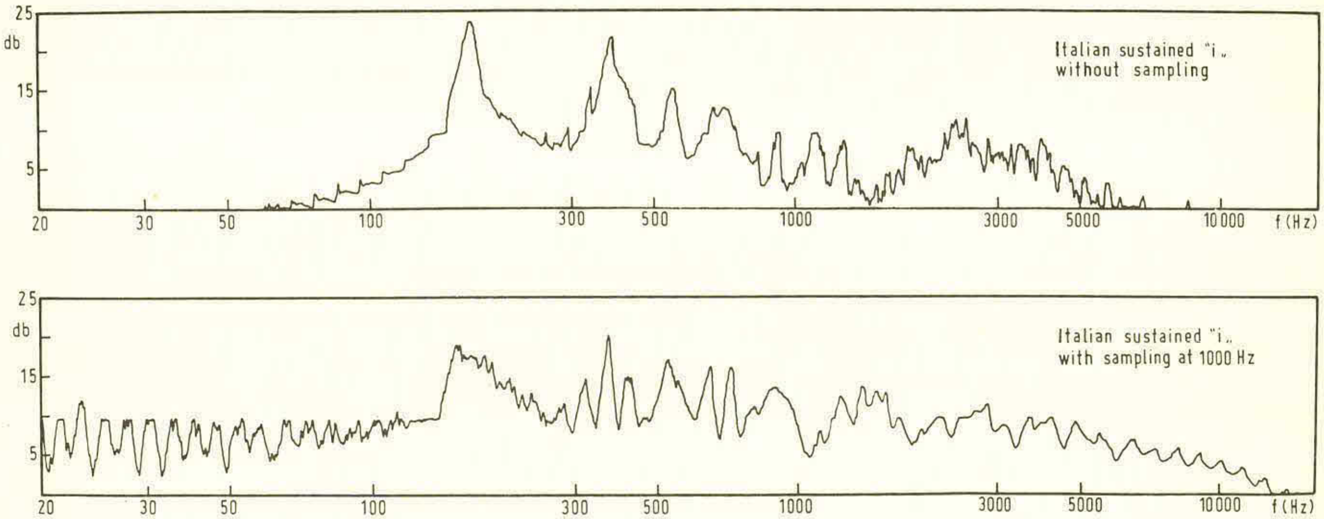


Fig. 8. Spectrum of Italian sustained "i".

categories:

a) Subjects with good audiograms (hard of hearing) and trained to the use of conventional prosthesis. They refuse, in many cases quite firmly, to use a sampled signal prosthesis. These subjects declare that they have already obtained some advantages by using different apparatus.

b) Subjects with bad audiograms and/or little training in the use of conventional equipment have shown to prefer the sampling prosthesis.

Satisfactory results have been obtained among the latter. Since the first sessions it has been possible to make most of these subjects distinguish some different words; the same results were not achieved by means of conventional equipment. Many of the people in the said conditions submitted themselves to a training over a period of many months and in some cases over an entire year.

Some of the subjects which endured pursuing such training are now capable of practically distinguishing any word and engaging themselves in simple talking provided the teacher pronounces every word lowly and pauses to stress word separation and repeats a few times the words possibly not understood. These subjects manage also to perceive (and therefore to repeat) words of which they are not aware and sentences without logical meaning.

The sampled signal prosthesis is provided with a knob by which, on the free choice of the subject, the sampling frequency may be varied between 1,500 and 5,000 Hz. The subjects have always been allowed to determine empirically the sampling frequency most convenient for them. In most cases, the selected frequency has ranged between 2,000 and 3,000 Hz. In this connection it is observed that the sampling would have brought about a minor talk distortion, if the frequency was 5,000 Hz, but at such a frequency its intrinsical advantage would have been reduced.

The fact that the subjects have selected, as an optimal sampling frequency, a value below 5,000 Hz proves the advantage of the method. As a matter of fact, in the hypothesis of the method being without benefit, a frequency of 5,000 Hz would have been selected according to which speech would have practically undergone a simple amplification with very modest distortion. In this connection, an interesting phenomenon has been noted, i.e., the subject's tendency to select higher and higher sampling frequencies as training progresses. Some of the children, after a period during which they have been trained to the use of sampled signal prosthesis, also profit by conventional prosthesis. A tentative explanation of this is shown in (6) and (7).

References

1. Biondi, F. and Biondi, L., "The Sampling of Sounds as a New Means of Making Speech Intelligible to the Profoundly Deaf," 1968, Alta Frequenza, pg. 728/739.

2. Biondi, F. and Schmid, R., "Mathematical Models and Prostheses for Sense Organs," on: Theory and Applications of Variable Structure Systems, (Ed. R.R. Mohler-A. Ruberti), 1972, Academic Press.

3. Kiang, N.Y.S., Watanabe, T., Thomas, E.C. and Clark, L.F., Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, 1965, M.I.T. Press.

4. Siebert, W. M., "Stimulus Transformations in the Peripheral Auditory System," in: Recognizing Patterns, Studies in Living and Automatic Systems, (Ed. P.A. Kolers - M. Eden), 1968, M.I.T. Press.

5. Thomas, I.B. and Sparks, D.W., "Discrimination of Filtered-Clipped Speech by Hearing Impaired Subjects," 1971, J.A.S.A., pg. 1881/1887.

6. Biondi, E. and Biondi, L., "Risultati di alcune sperimentazioni della protesi a segnali campionati mediante prove di comprensione del parlato da parte di sordi profondi," Rel. Int. Lab. Contr. Aut. 70-5, Ist. Elettronica, Politecnico di Milano.

7. Biondi, F., "Uso di modelli matematici in problemi Bioingegneristici per sistemi sensoriali menomati," Atti Congr. Cibernetica, 1971, Casciana Terme.

THE VSTA: AN APPROACH TO THE SPEECH TRAINING INSTRUMENTATION PROBLEM

by

L. C. Stewart
Robert A. Houde
Wilbur D. Larkin
Center for Communications Research
50 West Main Street
Rochester, New York 14614

Summary. The problem of teaching speech to the deaf would appear to be an obvious application for speech signal analysis techniques. Immediate feedback of speech performance might be expected to mediate the learning of speech rather directly. Over the past few decades a wide range of electronics instruments have been developed as speech training aids, yet none have ever achieved widespread use. This perplexing problem was examined with the objective of identifying the complex needs of the speech training situation, and the Visual Speech Training Aid (VSTA) was developed to satisfy these needs. The system utilizes digital storage and display techniques to present speech parameters as continuous traces on a television screen. The parameters are displayed in "flow mode" or strip chart mode. Two parameters from the same speech signal (for example pitch and intensity) may be recorded and displayed simultaneously, and the last few seconds of a display may be stored indefinitely. The target trace of a teacher can be displayed simultaneously with the attempt of the student. Teacher and student parameters are separated by using black and white traces on a grey background. Both teacher and student traces can be independently moved about on the screen and overlaid for direct comparison. Selecting parameters, recording the display and moving the display is accomplished through a convenient set of pushbuttons. The instrument is presently equipped to display intensity, pitch, the principal frequency of voiceless sounds and nasality.

Introduction

Speech is perhaps the most complex motor behavior carried out by man. It requires the performance of a set of coordinated gestures and activities with great precision. In the normal hearing child these complex skills develop naturally as a language - a code held in common with the other persons within the culture.

Early profound deafness (prelingual deafness) deprives a child of the means of acquiring speech skills and language naturally. Speech development in the deaf child is accomplished only through "training" procedures, that is, procedures which explicitly take into account the sensory deficiency preventing normal speech development.

The inability to produce speech is but one aspect of a set of interrelated communication handicaps deriving from early deafness. It can be argued that the language deficiency resulting from early communication deprivation constitutes an even greater handicap to normal intellectual and social development than either the inability to produce speech or the underlying auditory deficiency. The emphasis on speech production exhibited in this paper indicates to some degree the importance which we attach to this aspect of the problem, however, not to the exclusion of our concern with the other aspects of this complex situation.

Speech Training for the Deaf

Training techniques range from strict "natural" approaches to those which are extremely analytical. A natural approach attempts to expose the child to a wide range of acoustic and speech communication experiences. The natural approach is based on the notion that speech development requires a full range of natural communications experiences (which generally defy systematization) and can take place if the child's communication environment is suitably modified to compensate for the auditory deficiency by powerful acoustic amplification and by emphasizing clear face-to-face situations which encourage lipreading. An analytical approach involves the systematic identification of deficiencies in speech production and the use of specific training routines aimed at overcoming these deficiencies. Most speech training programs appear to be a blend of these two approaches, combining formal diagnostic and practice procedures with the encouragement of informal natural communication experiences.

The effectiveness of conventional speech training is quite difficult to ascertain; there is very little hard data on this question. It is quite well established, however, that successful speech training requires very large amounts of individualized work by skilled teachers. (inordinately large amounts in the view of non-oralists) At some risk of oversimplification we suggest that the principal reason why most prelingually deaf individuals do not achieve intelligible speech is that the quantity and quality of individual training required is greater than they or society are prepared to undertake.

Speech Training Aids

The obvious applicability of speech signal analysis and display technology to the speech training problem has resulted in a large number of sensory aid developments over the last decade (1-6). These developments range from a large computer based system by Bolt Baranek & Newman Inc. (BBN), (6), to the relatively simple single feature displays of the "Swedish" instruments (7). The system under development by BBN consists of a general purpose computer (PDP/8E) which accepts inputs from a spectrum analyzer, a pitch analyzer or a nasality indicator and which provides output displays on a cathode ray tube. (tactile output

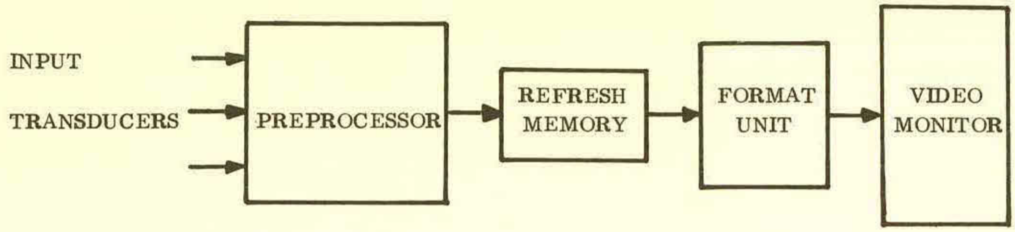


Figure 1. Block diagram of the VSTA.

displays are also planned for). This computer driven display constitutes a powerful experimental tool for exploring a large number of processing and display techniques. A major factor in determining the general applicability of this system to the problem will be its cost per student served. At the other extreme the relatively simple inexpensive instruments such as the Swedish "S" meter were developed on the rationale that simple low cost instruments naturally get wider distribution and therefore provide a greater total solution to the problem.

The Visual Speech Training Aid described in this paper is an attempt to achieve a compromise between the flexibility and completeness of the expensive computer based system, and the narrow but inexpensive simple instruments. The dominant constraint guiding this development was a projected instrument cost comparable to that of a medium range audiometer. The principal features of the VSTA which we found could be achieved under this constraint are the following:*

1. Speech Parameter modularity. The system displays time varying speech parameters visually. The parameters presently provided are pitch, intensity, nasality, voiced/voicelessness and frequency centroid of voicelessness. The modular design allows the addition of other time varying parameters or the removal of present parameters.

2. Television - storage display. The time varying speech parameters are displayed through a television refresh memory which provides flexibility in both the display format and the choice of display equipment.

3. Teacher-student comparison. The storage display allows a teacher "model" to be stored on screen during any number of student attempts at matching the model. The parameter traces may be moved about on screen for direct teacher student comparison.

4. Flow-mode. The time varying speech parameter traces are displayed in flow-mode or Times Square mode, being "written" near the right edge of the screen and flowing to the left (as if on a strip of paper) as time passes. A range of time bases are provided.

5. Ease of Operation. Both the post-secondary school deaf students and the speech therapists have learned to operate the VSTA by observing a few trials, that is, without formal training of any kind.

*The identification of objectives for this instrument was carried out in close cooperation with the staff of the Communication Center of the National Technical Institute for the Deaf.

System Description

Figure 1 is a block diagram of the VSTA and the system in use is shown in Figure 2. The basic components of the system are: a conventional video monitor, an electronic digital refresh memory, a preprocessor unit, a format unit, and various input transducers. The preprocessor contains preamplifiers for the input transducers and speech parameter circuits. Parameters currently in use include intensity, fundamental frequency Fo, a voice/voiceless decision, fricative centroid, and nasality. Nasality is derived from an accelerometer attached to the nose while other parameters originate from a dynamic microphone. The preprocessor simultaneously converts two parameters into digital form for storage in the memory. The digital refresh memory is organized into four sections. Each section can store a two second portion of a speech parameter, allowing both a teacher and a student to store pitch and intensity of their utterances.

Pushbuttons are provided for selecting up to two parameters at a time for processing and storage. Other pushbuttons initiate storage of parameters or the horizontal and vertical positioning of parameters on the monitor after they are in storage. The format unit converts the parameters in storage into a raster format for presentation on a video monitor. Each of the four

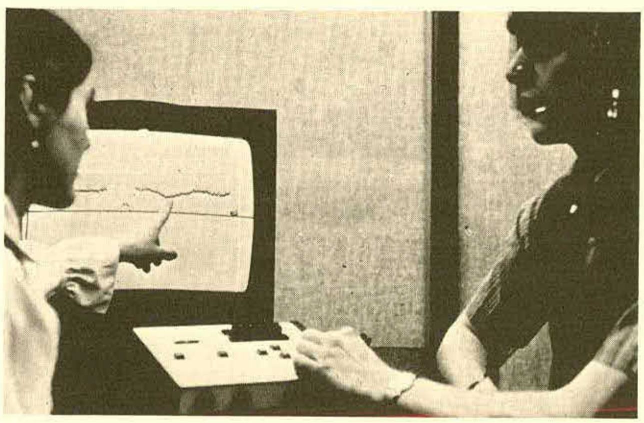


Figure 2. The VSTA in use.

parameters in storage appears as a narrow trace with time the abscissa and parameter value the ordinate. The screen is divided into top and bottom halves. A teacher's and a student's pitch traces, for example, may be overlayed on the top half while their corresponding intensity traces are overlayed on the bottom. Teacher and student traces are distinguished by using black and white traces on a gray background.

Parameters

The VSTA is basically a flexible visual display system which can display any speech parameter as long as it can be represented by a time varying voltage. The system is modular, allowing parameters to be added or changed at any time. The parameter circuits convert the acoustic or other transducer input signal into an analog voltage proportional to the parameter. Analog voltages are bandlimted to about 50 Hz. Parameter selection switches allow two parameters to operate at a time, one on the top half of the video screen, one on the bottom. Two analog-to-digital converters sample and digitize the parameters at 8.3 ms. intervals.

Fundamental frequency Fo is derived using a conventional configuration of lowpass filter to isolate the fundamental frequency and axis crossing meter to convert frequency to voltage. Two selections for cutoff frequency of the lowpass filter are provided. The higher position works well for females, children, and most males. Only the lower pitched males require the lower cutoff frequency. A range of 70 to 550 Hz. on a logarithmic scale was decided upon. This type of Fo tracker requires a good microphone. Several dynamic microphones have proven adequate. An ElectroVoice RE-51 boom microphone was particularly convenient and its close positioning to the mouth provided good signal to noise ratio.

Intensity is derived using a full wave rectifier and 50 Hz. smoothing filter. A rising preemphasis of 6 dB/oct. up to 3000 Hz. improves the presentation of fricatives. Dynamic range was limited to 30 dB. on a logarithmic scale for various reasons. Intensity has turned out to be a base parameter which is most often used with other parameters. It is useful therefore to keep the signal level for other parameters within reasonable limits and thereby ensure better operation of other parameter circuits.

The voice/voiceless parameter is intended mainly to indicate the presence of /s/ and /sh/. Fundamental frequency and fricative centroid parameter are displayed as one trace on the screen, with voice/voiceless parameter making a mutually exclusive decision. Operation of this parameter is based on the energy balance in two bands, one below 900 Hz. and one above 3500 Hz.

The fricative centroid parameter gives an indication of the frequency at which most of the energy is clustered in the fricative. It is most useful in distinguishing between /s/ and /sh/. A convenient arrangement for displaying fundamental frequency and the fricative centroid of voiceless sounds has evolved. During voicing Fo

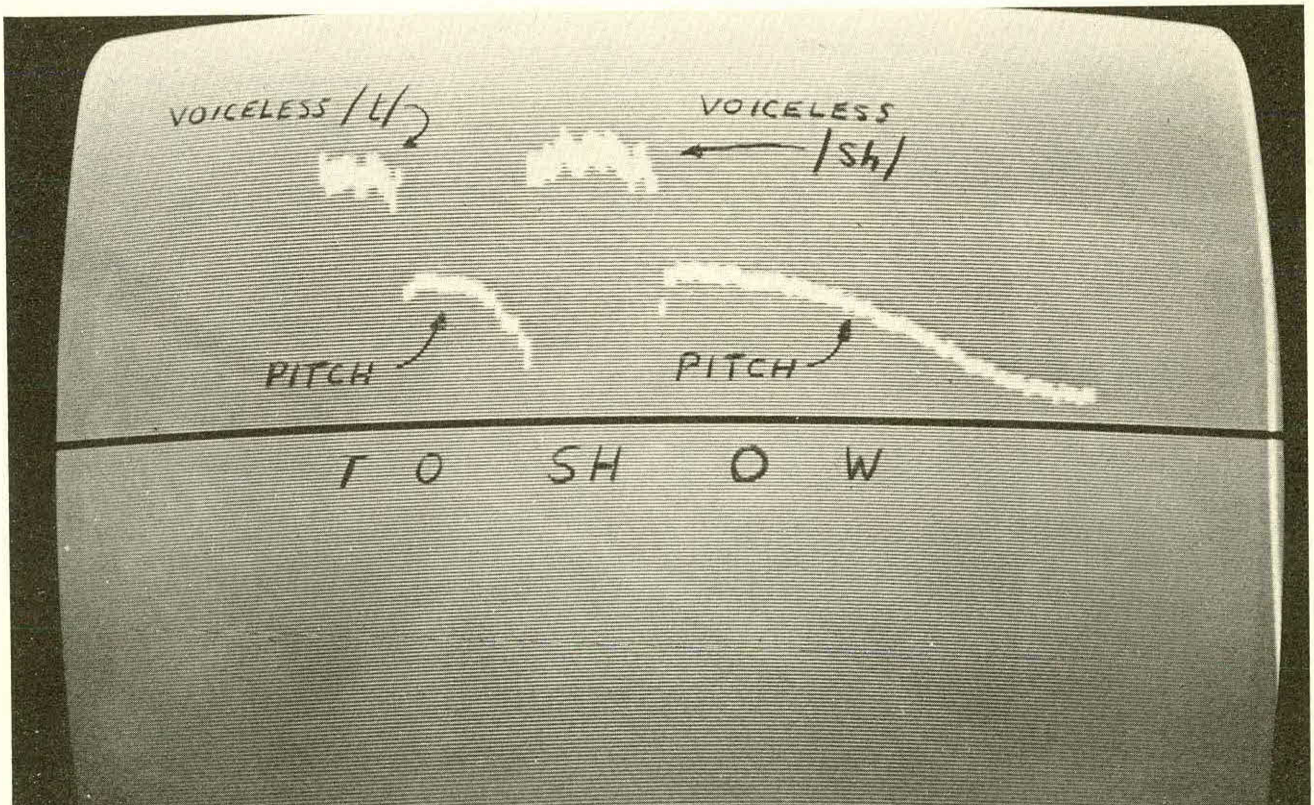


Figure 3. Video screen showing VSTA trace.

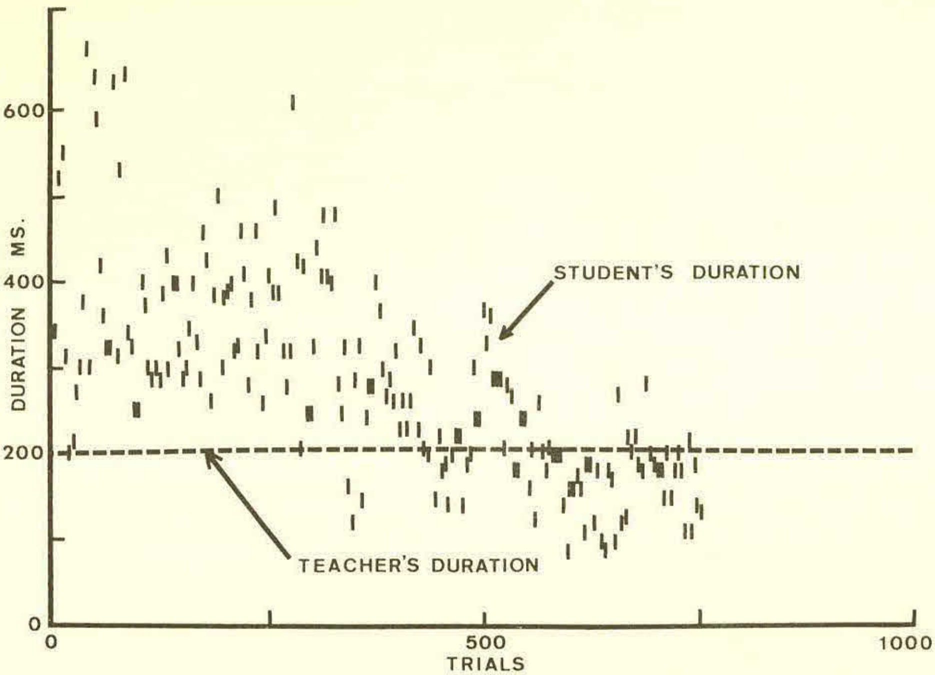


Figure 4. Performance record of one student.

appears on the screen as a narrow (4 TV lines in width) trace. Voiceless sounds cause the trace to be broadened to 8 TV lines and a different frequency scale is used for displaying the fricative centroid. Figure 3 shows a trace with pitch and voiceless portions. Most of the time Fo occupies the lower half of the scale (except for very high pitch). Fricative centroid on the other hand usually occupies the upper half scale. The combination of position and trace width makes distinguishing between the two quite easy. Voice/voiceless boundaries are likewise distinctively shown in the intensity trace by broadening the trace during voiceless intervals.

The nasality parameter attempts to show the relative degree of acoustic coupling through the velum by comparing the intensity of the vibrations on the nose surface to the intensity radiating from the mouth. A small light accelerometer (Bolt, Beranek Newman Model 501) is attached to the nose with two sided adhesive tape. The signal obtained is amplified and its intensity is measured by a circuit identical to the oral intensity circuit. The ratio of nasal intensity to oral intensity is derived and this signal is displayed on a logarithmic scale. Ratioing tends to make the nasal parameter independent of laryngeal intensity changes. The trace on the screen is at the low end of the scale for non-nasalized sounds. For semi-nasals it is about midscreen while strong nasalizing moves the trace to full scale.

Refresh Memory

An electronic digital memory stores the parameter samples and reads the values out at a fast rate for refreshing a video monitor. A digital refresh memory in conjunction with a video monitor provides considerable flexibility in both the recording of speech parameters and in subsequent viewing of the parameter traces: (1) the recorded parameters can be viewed indefinitely without deterioration; (2) the parameter traces can be moved about on the screen for overlay or comparison after recording is completed; (3) teacher and student traces can be erased separately, allowing a teacher's model to remain during repeated attempts by the student; (4) the refresh memory acts like a two second endless recording loop, allowing recording to continue indefinitely until a suitable two second portion has been obtained

During recording, the two second portion in storage also appears on the screen. The new speech appears on the right edge of the screen and progresses to the left and off the screen in what has become known as a "flow mode". When recording ceases, the part currently on the screen is saved in storage. It can be moved to the left by a horizontal positioning switch and as part of the trace moves off screen it reappears on the right edge, nothing being lost.

Format Unit

The four parameters in storage are converted into a TV raster format in the format unit. The screen is divided into top and bottom halves. Pitch, fricative centroid, or nasality appear on the upper screen while intensity occupies the lower screen. Teacher and student traces are easily distinguished by the use of black and white traces on a gray background. In the event of overlap the white trace dominates. The upper and lower screen each have 112 lines of TV resolution, adequate for any speech parameters. The traces are normally four TV lines in thickness to make them move visible. During unvoiced intervals the traces are broadened to 8 TV lines. The traces can be moved up or down on the screen by means of vertical position switches. When part of a trace reaches the top or bottom edge of its half screen field, it disappears "behind a mask" so that it in no way interferes with the other half screen field.

Video Monitor

A video monitor was selected as a display medium because: (1) monitors are becoming the standard medium in education; (2) a wide range of screen sizes is available (including large screens); (3) the combination of a digital refresh memory and monitor provides great flexibility in display. Although not used in the VSTA program thus far, any number of video monitors could be connected to a VSTA for special classroom situations.

Evaluation

A systematic exploration of the use of the VSTA is currently under way with profoundly deaf post-secondary school students. Final results of this evaluation are not yet available, however, the intermediate results shown in Figure 4 serve to illustrate a number of important factors in this problem area. Figure 4 summarizes the performance record of a student attempting to develop the production of a short (200 ms.) unstressed initial syllable in the phrase "to show". The pitch/voiceless trace of a satisfactory production of this phrase was shown previously in Figure 3. The duration of the student's initial attempts were too long and varied widely. After approximately 750 trials his performance exhibited a much tighter distribution which centered about the performance objective of 200 ms. During this training the student was essentially operating in a self instruction mode. The student was at ease with the process, highly motivated, and was able to accomplish successive trials at a very high rate, often approaching one trial per second. When we consider the large number of trials necessary to achieve this simple skill, it is seen that trial rate is an important factor in speech training. In this case the high trial rate is attributable to the flow mode display. Rather than stop and position each trial for comparison to the model, the students in this study adopted the tactic of comparing their attempts to the model "on the fly" which the flow mode makes possible.

It must be pointed out that other speech skills are not as responsive to training as was syllable duration. Preliminary results indicate that intonation is particularly resistant to training, as was observed in a recent study by Boothroyd (8).

The VSTA appears to be well suited to drill and practice procedures, while this is the very aspect of speech training for which the speech therapist is intellectually least well suited. Thus the promise of speech training instrumentation may be realized by complementing the diagnostic capabilities of the therapist with instrumentally aided drill and practice procedures. Particularly promising is the amenability of these procedures to monitoring by para-professionals and to self instruction.

References

1. W. Pronovost, "Developments in Visual displays of speech information", Volta Review, 1967, 69, 365-373.

2. J. M. Pickett, "Recent research on speech-analyzing aids for the deaf", IEEE Transactions on Audio and Electroacoustics, 1968, AU-16, 227-234.

3. J. M. Pickett, Ed., "Proc. conf. on speech analyzing aids for the deaf", Amer. Ann. Deaf, 1968, Vol 113, 116-330.

4. H. Levitt, "Speech processing aids for the deaf: An overview", IEEE Transactions on Audio and Electroacoustics, 1973, AU-21, 269-273.

5. H. Levitt, P. W. Nye, Ed., "Proc. conf. on sensory training aids for the hearing impaired", November 15-17, 1970, Committee on the Interplay of Engineering with Biology and Medicine, National Academy of Engineering.

6. R. S. Nickerson, K. N. Stevens, "Teaching speech to the deaf: can a computer help?" Proc. Nat. Conv. ACM, August, 1972, 240-250.

7. A. Risberg, "Visual aids for speech correction", Am. Ann. Deaf, 113, 178-194.

8. A. Boothroyd, "Some experiments on the control of voice in the profoundly deaf using a pitch extractor and storage oscilloscope display", IEEE Transactions on Audio and Electroacoustics, 1973, AU-21, 274-278.

A COMPUTER BASED VISUAL AID FOR THE DEAF*

by

Stanley C. Person and Ian B. Thomas**
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts 01002

Summary. A computerized visual display is being developed as a research tool for articulation training in the profoundly deaf. Efforts so far have been concentrated on the extraction of accurate formant information from the speech of children. This represents a significantly greater challenge than the extraction of such information from the voices of adults particularly in virtual real time, yet is essential if articulation training using visual displays is to become available at as early an age as possible. A mixed frequency and time domain analysis is performed primarily by the computer. Results are displayed in virtual real time on an oscilloscope display or are available in hard copy if desired. Speech is filtered into bands corresponding to the ranges of the first two formant frequencies (F1 and F2). The resulting waveforms are then analyzed by the computer for determination of the formant frequencies. Displays available include plots of F1 versus F2, or of these formants as explicit functions of time.

Introduction

The ability to learn to produce articulate speech depends on the integrity of a feedback path which in normal circumstances includes hearing. For the profoundly deaf, the customary feedback path has been essentially blocked by virtue of the profound hearing loss so that these individuals are forced to use some other means of obtaining feedback as to the correctness of the speech sounds they generate.

Until recently, closing of the feedback loop has required the presence and the cooperation of another individual with satisfactory hearing. It has also required the acquisition of other communication skills, primarily through the visual sense (lip reading, reading and writing), although considerable use has also been made of the tactual sense in teaching situations (feeling for nasalization, phonation). The shortcomings of these substitute feedback paths are most dramatically realized by simply listening to the speech of profoundly deaf individuals, including those more fortunate ones who have spent ten years or more in special schools for the deaf.

In response to the need for an improved feedback path, communication engineers have attempted to present "appropriate" feedback to the deaf in real time. Indeed, it is the promise of real time feedback which has been the most prized feature of technological aids for the deaf. The difficulty arises in choosing the most appropriate parameters to present to the deaf individuals given the fact that the visual and tactual senses have ultimate limitations on the amount of information which they can successfully process within a given time.

A very important consideration in the design of visual aids for the deaf is the age of the profoundly deaf individual to be assisted. It is clearly recognized that language acquisition is most easily and efficiently accomplished in the earliest years of life. Thus it pays to concentrate training for the congenitally deaf most heavily in these early years. This in turn implies that the output of the device used to provide feedback should be readily interpreted by very young children (e.g., three-year olds). Moreover, it must be capable of extracting the appropriate data from the voices of very young children. With the usual methods of speech analysis, this last consideration presents a significant problem which has received very little attention as yet. Because of the higher pitches of children's voices, analysis through frequency domain techniques is much less accurate than it is in the case of adult voices. It is this problem of the analysis of children's voices with which this work is primarily concerned.

Before considering the question of how to analyze adequately the voices of children a decision must be made as to which parameters to extract and in what manner these parameters should be presented to the child. Considerable guidance is available on this question from work already accomplished here and in other laboratories (1-6). In a study completed in 1970 (4), very encouraging results were obtained in articulation training through the use of a real time display of first versus second formant frequency. The arguments for use of this type of display are clearly detailed in that study and will not be repeated here. The study subjects were all deaf adult males so that the device performed adequately in the extraction of formant data using a bank-of-filters (frequency analysis) approach. When the device was tested with children's voices the limitations of the analysis method became obvious so that our attention then turned to other ways of obtaining accurate formant data from all voices. We decided to use a simple time-domain

---

* Please address all correspondence to Professor Ian B. Thomas, Electrical and Computer Engineering Department, University of Massachusetts, Amherst, Massachusetts 01002

** Also in the Department of Communication Studies.

analysis technique which we have found yields good results in speech recognition and formant tracking using adult voices (7). Our most recent work of applying this analysis method to the voices of children is presented in this study.

Theory

For purposes of discussion let us assume that the waveform corresponding to a voiced (phonated) speech sound is periodic. Its frequency spectrum will then consist of a set of discrete lines at the fundamental (glottal) frequency and at each of its harmonics. Formant frequencies can be estimated from such a spectrum by looking for local maxima in the envelope drawn over the discrete frequency spectrum or, somewhat more directly, by saying that the formant frequencies can be estimated as the frequencies of the locally tallest lines in the spectrum. If the glottal frequency (pitch) of the voice is, for example, 100 Hz (corresponding to the pitch of a typical male voice), then there will be an effective sampling of the spectrum envelope at all the integral multiples of 100 Hz. Given the typical frequency ranges, bandwidth, and difference limens of the formants this sampling of the spectrum is adequate. It defines formant frequencies with absolute accuracy of $\pm$ 50 Hz ($\pm$ fg/2 where fg is the glottal frequency).

Suppose now that the glottal frequency is changed to 400 Hz without any change in the configuration or size of the vocal tract (and thus the spectrum envelope). Now the lines of the spectrum begin at 400 Hz (the fundamental) and are found at all integral multiples of this frequency. Again, if we use the frequencies of the locally tallest lines in the spectrum as estimators of the formant frequencies, we encounter an absolute error of $\pm$ 200 Hz, which is unsatisfactory for distinguishing speech sounds from one another. This case represents a slightly exaggerated view of the difficulty of estimating formant frequencies in a child's voice by spectral analysis techniques. Because of the smaller dimensions of their vocal tracts, children's formant frequencies are from 20% to 30% higher than those for adult males but this slight increase in the formant frequencies hardly compensates for the factor of 3 or more increase in the glottal frequency.

Let us take a somewhat different theoretical approach and assume that each glottal period of a voiced speech waveform consists of a sum of damped sinusoidal components. There is one damped sinusoid corresponding to each formant and the "ringing" rate of each damped sinusoid corresponds to the formant frequency. Thus if one can isolate each of the separate damped sinusoidal components through broad band filtering into the expected formant frequency ranges, it should be possible to obtain an accurate estimate of the formant frequency by measuring the duration between successive zero axis crossings of the filtered waveform. Theoretically, this method is as applicable to the speech of children as to that of adults. For children's speech, the basic periods (glottal periods) of the speech waveform will be shorter than those for adults (because of the higher pitch) thus giving fewer axis crossing intervals per period with which to estimate the formant frequency. This factor becomes crucial only when there is considerable noise interference and the accuracy of the formant frequency estimate depends on the number of axis crossing intervals over which an average can be taken per glottal period. As mentioned earlier, this analysis technique has been tested successfully for both formant tracking and speech recognition work with considerable sources using adult voices (7-9).

If the speech waveform is filtered using a bandpass filter adjusted to pass the frequency range in which the formant is expected the resultant waveform will ideally be a sequence of damped sinusoids whose frequencies correspond to the desired formant frequencies. Actually, the filtering is configured to make the formant component of interest the one of the largest amplitude in the frequency region of concern. However, perfect separation is not achieved since the frequency ranges of each formant overlap. In order to minimize the overlap effect the cut-off frequencies are selected so that the filter characteristic exhibits a falling slope in the region of overlap.

An additional and more serious problem that arises with increasing pitch is the greater variability of axis-crossing intervals. If it is noted that each glottal pulse is not phase locked to previous pulses or to the damped sinusoidal response, then it should become clear that as the pitch increases (increasing frequency of vocal tract excitation) there will be an increasing number of long (or short) axis-crossing intervals. Each new excitation can come at any point in the waveform, thus even though the resonant frequency is the same, there may be long or short intervals present.

Another source of erroneous data is noise in the surrounding area. When the primary stimulus has a low amplitude, background noise or speech can effect the data and the resultant formant track. There is little that can be done for this situation except to try and maintain a rather quiet area when processing in real-time from a microphone. From experience in operating the system, the background sound level has to get very high before the formant values are affected to such an extent that the displays are unusable.

The solution to the problem of variability in the axis-crossing durations lies in smoothing the data by rejecting data values that fall outside of the expected frequency range. Also, the remaining data is averaged over a time interval chosen to be long enough to provide a meaningful average. These steps are taken to remove the variations in the axis-crossing intervals due to the high pitch. It has been found that the reciprocal of the resulting axis-crossing interval yields a good approximation to twice the formant frequency for the averaging interval. For example, suppose an F1 (first formant) track is in progress. The range for F1 is approximately 300 Hz to 1 KHz corresponding to axis-crossing intervals of greater than 0.5 msec but less than 1.66 msec. Any interval outside of this range would be discarded by the algorithm. The rest of the data would be averaged for some time interval, and the formant frequency for that interval would be determined by taking one-half the reciprocal of the resulting axis-crossing interval.

The choice of the length of the sampling interval involves a trade-off. If a short interval is chosen (say 5 to 10 msec) so that it is possible to follow rapid formant transitions, then the number of axis-crossing intervals will be small, increasing the effects of sample variability as previously mentioned. Similarly, a long sampling time decreases the ability to follow rapid and subtle formant transitions, but increases the immunity to noise and sample variability. It has been found that a reasonable range of intervals is 1/30 to 1/60 second. The number of axis-crossings for an interval in this range is sufficiently large so as to provide a statistically meaningful average. Assuming that the signal frequency is 1 KHz, there will be 500 microseconds between axis-crossings. That corresponds to at least 64 intervals during a 1/30 second sampling interval.

Summarizing the method, assume that a voiced speech signal consists of a sum of damped sinusoids at the formant frequencies. Then, if the waveform is bandpass filtered to isolate a single formant, the measured axis-crossing intervals of the resulting signal will yield a good estimate of the reciprocal of twice the formant frequency. In order to follow the time transitions of the formants the following algorithm is proposed:

Let

T_i be the i^{th} axis-crossing interval in the sampling period.

L be the lower bound for allowable intervals

H be the upper bound for allowable intervals

N be the number of intervals in the averaging period such that $L<T_i<H$ (smoothing operation)

Then, calculate the average axis-crossing interval T_{ave}:

$$T_{ave} = (1/N) \sum_{i=1}^{N} T_i$$

The formant frequency for the sampling interval being considered is:

$$f = 1/(2\ T_{ave})$$

This algorithm is obviously adaptable to real-time computer techniques. Also, it will be seen that very little auxiliary hardware is needed when compared to the standard frequency domain approach using a bank of hardware filters.

Although this algorithm has been designed by assuming that the speech sample is voiced, formant tracking is possible for other classes of speech sounds. Consider first the fricative consonants. The fricatives are produced by a noise excitation of the vocal tract at a constriction. There are both voiced and unvoiced fricatives. For the unvoiced fricatives the axis-crossing intervals of the bandpass filtered waveform give a poor estimate of the formant frequencies since most of the energy in these sounds is concentrated above 4 KHz. For the purpose of articulation training this problem can be resolved by noting that a given fricative articulatory configuration can be excited with or without voicing (these voiced-unvoiced pairs with identical articulation are called cognates). The addition of voicing provides sufficient low frequency energy so that the fricative may be treated the same as a vowel and similar accuracy may be expected. The voiced-unvoiced pairs are:

Voiced	Unvoiced
/v/-vote	/f/-for
/ð/-then	/θ/-thin
/z/-zoo	/s/-see
/ʒ/-azure	/ʃ/-she

Another class of speech sounds different from the voiced vowels originally considered are the nasal consonants. These sounds are voiced, but due to participation of the nasal tract, extra poles and zeros are introduced into the transfer function of the vocal tract (10). Due to pole-zero cancellations poor results follow. At the present time a satisfactory solution to the problem of formant tracking nasal sounds has not been incorporated in this algorithm. It should be noted that the values determined by the axis-crossing interval method will correspond to any significant resonances found in the regions of the frequency spectrum being considered.

The glides and dipthongs, along with the semivowels and affricates do not present any theoretical difficulty other than those discussed for the voiced sounds. All of these phonemes are voiced and have no effective nasal interaction, hence the speech waveform produced will not violate the underlying assumptions of the formant tracking algorithm being considered.

The stop consonants have not been included in the testing of this algorithm. The reason for this omission is that it was found to be difficult to conduct an effective accuracy test for these sounds with the established hardware and software. The cause for this difficulty is the short duration of the stop burst which brings some timing problems into play. It was found to be very difficult to accurately locate the burst in the vast volume of axis-crossing data that is output by the system. Solution of this problem is beyond the scope of the current research. Additionally, since transitions of the formants can be followed up to the point of the release of the burst and that it has been theorized that the stops are characterized by their transitions, it was felt that this concession was a small one.

Theoretically then, this algorithm will handle all voiced speech sounds except the nasals and stops where inaccuracies result. The formant structure of the fricatives can be followed with a special purpose filtering configuration that takes into account the high frequency energy concentration of these speech sounds. If the general purpose filter configuration is used, then the voiced equivalent should be used in articulation training as this provides additional energy in the F1 and F2 ranges.

System Description

Figure 1 is a block diagram of the system used to implement the axis-crossing interval formant tracking algorithm. The system can follow the transitions of any one of the first two formants in real-time and display it on an oscilloscope (Hewlett Packard X-Y Display) as well as provide numerical and graphical hardcopy. The data available on the line printer include individual axis-crossing intervals grouped in 16.6 msec time periods (1/60 second); computed formant frequencies for an averaging interval of either 16.6 msec or 33.2 msec tabulated along with a graphical representation of formant frequency versus time. All outputs are indexed so that the individual axis-crossings for any calculated formant can be readily located.

The speech waveform is obtained from either a microphone and amplifier or tape recorder output.

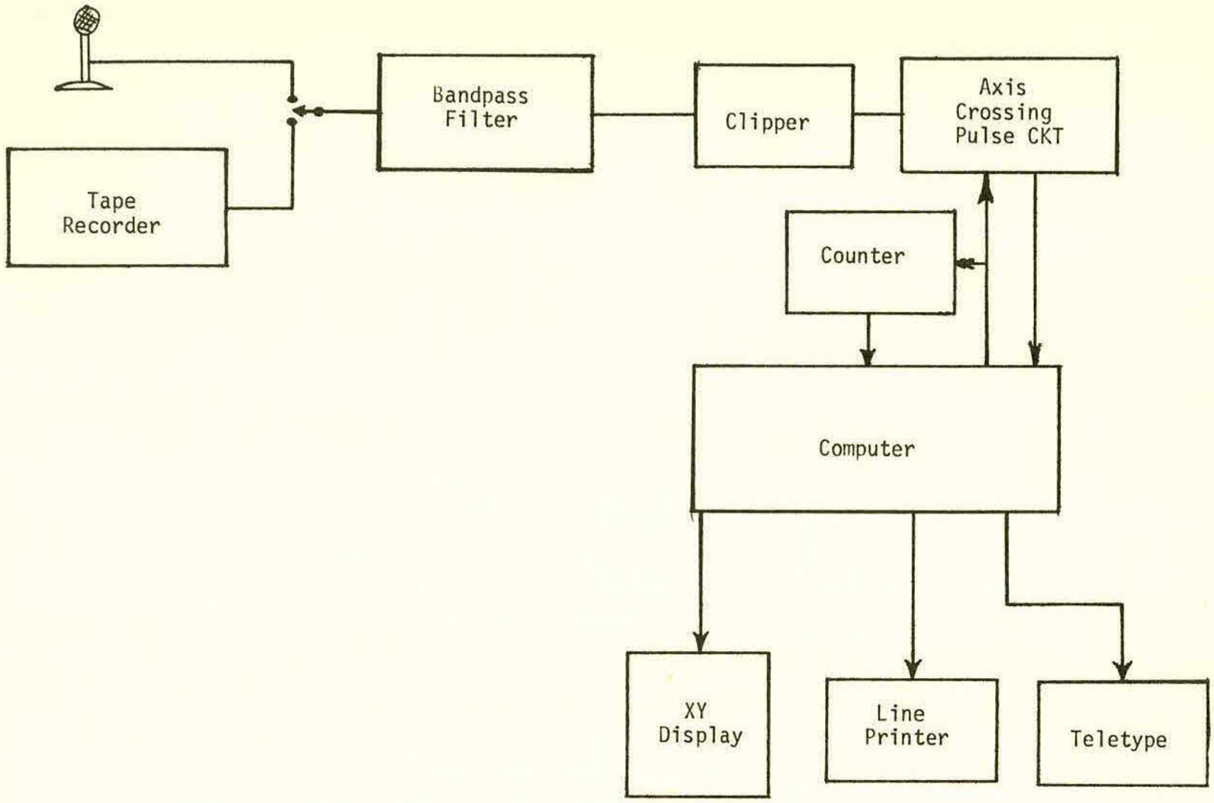


Fig. 1: System Block Diagram

The signal is then filtered by a bandpass filter adjusted to pass those frequencies in the region of the frequency spectrum that corresponds to the desired formant. The resultant waveform should be at least 50 millivolts in amplitude for reliable operation of the clipper. The filtered signal is infinitely clipped (amplitude of 5 volts), and sent to a logic circuit that produces a pulse each time the waveform changes polarity.

This pulse is sent to the PDP-11 computer, and upon receipt the computer interrogates a free-running counter (10 microseconds per count), obtaining the time interval since the last pulse. This interval is the time between axis-crossings of the filtered speech waveform which is the desired parameter. Immediately upon completion of interrogating the counter the computer generates a reset pulse that clears the counter and axis-crossing pulse circuitry, enabling receipt of subsequent axis-crossing pulses. This process is continued until a clock interrupt marks the end of an averaging period and the beginning of a new sample. The time interval between the start of an axis-crossing pulse and the generation of a reset pulse is a constant (about 8 microseconds), and therefore has no effect on the data for all practical purposes. The only time this delay could cause trouble would be if another axis-crossing came before the PDP-11 produced the reset pulse, but this would be a pathological case since it corresponds to a frequency that is out of the range to be considered (less than 3.5 KHz). That interval would be discarded by the algorithm even if it were input to the computer.

The choice of clipping the waveform was made solely on the basis of simplifying the chore of calculating axis-crossing intervals. It is feasible, in theory at least, to sample the waveform using an analog to digital converter and detect axis-crossings by sign changes of the value input to the computer from the A/D converter. The actual axis-crossing time could then be determined by a linear interpolation, and the interval could be found by subtracting the previous crossing time. Unfortunately, this calculation requires several divisions and multiplications so that the total time to calculate an axis-crossing interval is on the order of 100 microseconds. In order to maintain real-time response this calculation must be completed before the next axis-crossing occurs. The upper limit for the frequency that could be processed in real-time would be 5 KHz (or 1/200 $(\text{usec.})^{-1}$). This does not include the time for handling the individual samples from the A/D converter, processing interrupts, formatting, calculating the formant values from the data, or general overhead. The upper frequency limit when these factors are considered would be less than 2.5 KHz, and at level of activity there would be no room for expanding the system capabilities. Sampling the waveform using an A/D converter is clearly not an adequate solution to the problem of determining axis-crossing intervals in this case.

In order to reduce the computation overhead to a minimum, it was decided to input a measure of the axis-crossing intervals directly into the computer. Since the only parameter of the filtered speech that is of interest to this algorithm is the time interval between axis-crossings, the previously mentioned clipping operation is not only allowable, but it appears to be a logical choice since it strips the waveform of all "unnecessary" information. The axis-crossing times are maintained and the waveform is reduced to a form that is suitable for processing by a simple digital circuit.

The pulse train that corresponds to the clipped waveform is input to a logic circuit that produces a logic 0 to logic 1 transition at the Request A input of the General Purpose Digital Interface (DR-11A) of the PDP-11 computer. This transition alerts the computer that an axis-crossing has occurred in the speech waveform. The program causes the computer to interrogate the low-order byte (8 bits) of the DR-11A input word. This part of the input word is connected to the counter which has been incrementing at a 100 KHz rate since the last axis-crossing so it contains a count proportional to the current axis-crossing interval. After the value in the counter has been read by the computer, a reset pulse is output by complementing the DR-11A output word twice in succession. The reset pulse clears the counter and causes a logic 1 to 0 transition at the Request A input enabling further axis-crossing pulses to be received. This interplay is presented in Figure 2. A clock interrupt provided by the internal line-frequency clock informs the computer program of the necessary timing information for calculating the formants (i.e., provides the averaging interval time markers).

The speech processor was designed as a self contained unit with clipper, counter, axis-crossing

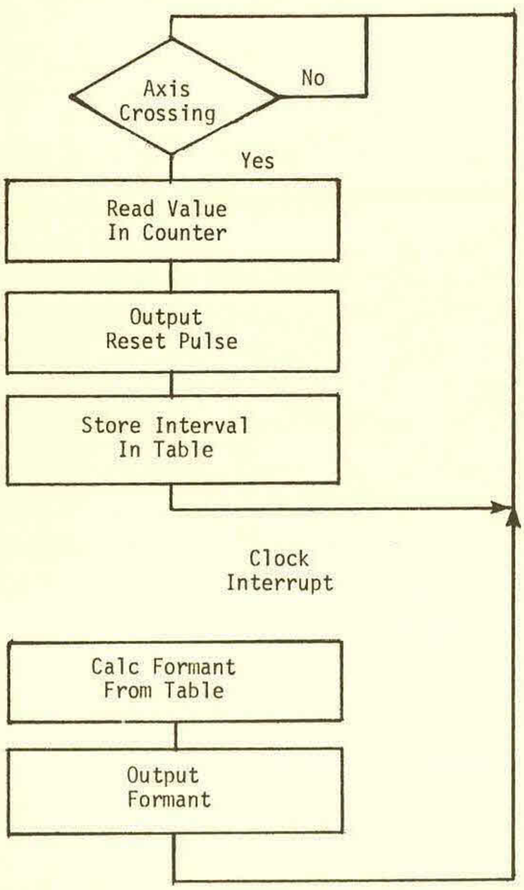


Fig. 2: Flow Diagram

pulse circuitry, clock and power supply mounted on a single chassis. Each functional component of the processor is mounted on a separate printed circuit board. Inputs and outputs from the unit are provided by BNC and banana plugs.

The counting rate is 10 microseconds per count so that axis-crossing times can be located to within that time span. This corresponds to an accuracy of 6% for a 3 KHz input or 2% for a 1 KHz input. The system was rigorously tested using sinusoidal inputs. The frequency of the input was measured using a calibrated counter and the results were compared with the computer calculations. The accuracy for tests was as determined above. The individual axis-crossing intervals output by the system were carefully scrutinized to verify consistent performance as would be expected with sinusoidal inputs. The task remains to show that the system does in fact measure the formant frequencies of a speech waveform. It has been demonstrated that the system is capable of determining the frequencies of sinusoidal inputs up to 5 KHz to an accuracy of better than 10%. It should be pointed out at this time that the difference limen for the first and second formants is on the order of 5%.

Results

The task remaining before us is to prove that the formant tracking system described actually does follow the time transitions of the resonances of the vocal tract for children's voices. It is also necessary to provide a quantitative measure of the formant tracking accuracy. A rigorous method of comparing system output, spectrograms, spectrogram sections, manual calculations and literature values is outlined below. This method demonstrates the fact that the formant frequencies are successfully isolated and tracked by the system. The vowels, fricatives and nasals are each tested, first as continuant sounds and then in normal speech.

For each speech sound tested, the subject (child) was to sustain the sound for about 2 to 3 seconds. After producing the sound in this continuant fashion, the subject was then asked to use the phoneme in a word. This process was repeated until each of the vowels, fricatives, nasals, glides and dipthongs were recorded for each of the three children.

The next step was to input each of the speech sounds into the formant tracking system. The system was set up to follow the first formant (300 Hz to 1 KHz). The individual axis-crossing intervals were output along with the computed formant frequencies and the graph of formant frequency versus time. The process was then repeated for the second formant (prefiltering set at 1.25 KHz and 3 KHz). This step provides the values of formant frequency as computed by the formant tracking system. This sequence of operations was continued for each phoneme produced by the subjects.

In order to accurately correlate the large volume of system output to the speech sample, the real-time formant frequency display was used. In this manner each phoneme could be accurately located in the printout and the formant value could then be determined. These values were tabulated and compared with literature values for F1 and F2.

The literature values are merely a collection of average resonance characteristics, so that a good comparison with the values output by the system is only mildly satisfying. In order to enhance confidence in the accuracy of the formant tracking system an additional step was taken, namely to analyze sections of the spectrograms of the phonemes. The spectrogram section is a display of spectral amplitude versus frequency at a particular point in time. By approximating the spectral envelope, an estimate of the resonances can be obtained from the display (with the limitations mentioned earlier for frequency analysis methods). It must be noted that this is only an estimate, but when considered with the tabulated characteristics a very good idea of the formant values can be obtained. The value obtained in this fashion is used as the standard from which the percentage error is measured.

As a final check, the prefiltered speech was input to a storage oscilloscope in order to visually check the waveforms. Samples of axis-crossing intervals were obtained by looking at the traces stored on the oscilloscope screen. A formant calculation was then made manually for the phoneme input to the scope. Naturally, this technique applies only to continuant speech, and since only a small number of axis-crossing intervals can be processed in this way the technique is subject to rather gross inaccuracies. However, in certain cases the resonance is particularly clear from the waveform and then this technique does give valuable information pertaining to accuracy. Taken along with the spectral analysis and the literature values a formant frequency for the particular speech sample can be arrived at with a good degree of confidence.

The results of these operations can be compared to the formant frequency determined by the system and a reasonable measure of accuracy can be derived from this data. If variations from the tabulated formant frequencies seem large, one must consider that the values are averages over a large number of subjects, and that the individual realizations can vary over a significant range.

The previous discussion has outlined the technique used to demonstrate that the formant tracking system does in fact follow the time transitions of the first and second resonances of the vocal tract. The following sections will discuss the observed results for each class of phonemes tested. Finally, some examples of formant tracking in sentences will be presented and discussed.

Vowels. The vowels are the first group of phonemes to be discussed. As explained in the test procedure, the vowels were voiced in a continuant fashion and then were used in a word (e.g., /i/ as in eve). Two evaluations of the system output were made for this group of phonemes since proof of system accuracy for the vowels is of great importance to the relevance of this method. First, the average system value for F1 and F2 for the three subjects was compared to formant values tabulated in the literature. The results of this comparison are tabulated in Table I. The largest deviations for F1 are less than 20% of the literature value and less than 15% for F2. In an attempt to find the reason for these discrepancies the values for a particular subject were compared to spectrogram sections of each phoneme.

Table I: Formant tracking results for vowels as compared to literature values.

Phoneme	Literature F1/F2	Ave. Syst. Output F1/F2
/i/	370/3200	380/3170
/I/	530/2730	575/2670
/ɛ/	690/2610	670/2350
/æ/	1010/2320	820/2380
/a/	1030/1370	900/1570
/ɔ/	680/1060	750/1170
/U/	560/1410	500/1370
/u/	430/1170	450/1070
/ʌ/	850/1590	800/1480

Table II: Formant tracking results for vowels as compared to spectrogram values.

Phoneme	Section Value F1/F2	System Output
/i/	480/3300	400/3300
/I/	500/2500	550/2600
/ɛ/	700/2300	700/2400
/æ/	800/2250	800/2250
/a/	900/1700	850/1600
/ɔ/	750/1200	700/1200
/U/	500/1200	500/1300
/u/	490/1100	400/1100
/ʌ/	700/1500	700/1500

The results in Table II show an improvement, with the largest deviation being 8.3% and many of the system F2 values were in complete agreement with the formants determined from analysis of the spectrogram sections. The average deviation for F2 from the section value was less than 2.5%. The results for F1 also showed an improvement, but not as dramatic as for F2.

The largest deviation of the system F1 value when compared to formant values determined from the spectrogram sections was less than 18.5%. The large errors came from the phonemes exhibiting low first formants (less than 500 Hz). These errors appear to have been caused by a baseline noise problem in the Kay Spectrum Analyzer since the phonemes exhibiting large F1 error as determined by this method show quite low error when the average values are compared as in Table I.

The formant isolation problem is clearest in this group of phonemes. In particular, note the elevated F2 values for /ɔ/ (all) when it is compared to the literature value. This is caused by the close proximity of F2 which has the effect of raising the average value determined for the first formant. Little can be done for this problem since the high F1 values of /a/ and /æ/ preclude lowering the upper cutoff frequency of the bandpass filter.

Fricatives. The spectral structure of the fricatives has been studied by Strevens (11) and Heinz and Stevens (12). The subjects appear to have been adult males. However, since the spectra of the fricatives are subject to such great variability, their data are used as a rough basis for comparison of the system output along with the spectrogram sections. These data values were found to be in good agreement with the spectral sections made of the actual speech segments analyzed, but it was impossible to find literature values for the voiced fricatives to be used as a basis for comparison. The voiced fricatives are made up of two excitation components: a hiss and a glottal excitation component. It is reasonable to assume that the acoustic characteristics of the hiss will correspond in most respects to those of the voiceless fricative. The basic difference in articulation lies in the fact that the air-flow is less for the voiced sounds than for the unvoiced since the airstream is interrupted by the vocal cords. Also, the high frequency energy will be less for the voiced sounds than the unvoiced. Nevertheless, without any other data, the same resonances will be assumed to hold true for each cognate pair of fricatives. Table III was constructed with these facts in mind.

It is not surprising that the first formant values determined by the formant tracking system do not agree with the literature values. The lowest resonance for any of the fricatives is 1500 Hz which is well out of the passband of the first formant filter (upper cutoff at 800 Hz). Investigation of the speech waveform filtered by the F1 filter did reveal a low level 500-600 Hz component for each of these phonemes. However, this could hardly be called a resonance since it was at least 30 dB down from the peak amplitude at major resonances. The point to be made from this observation is that while the system did not find the first formant, it did find the frequency of the low level signal that was present in the region of the frequency spectrum that was under investigation. For this reason the deviation percentages shown in Table III are figured using the formant that would be found in the normal F2 range. That is, the F2 system output corresponds to the expected F1 values.

With these assumptions, the formant tracking system locked onto a resonance of the voiced and unvoiced fricatives with an average deviation of 2.2%. It must be understood that system outputs for F2 are actually the first formant values. Accurate determination of the actual F2 values could be made with adjustments to the bandpass filter settings and some software modifications. The software changes are necessary to allow processing of the high frequency signals which would otherwise be smoothed out by the algorithm.

Summarizing the results for the fricatives, the formant tracking system accurately determines the first resonance of the voiced and unvoiced fricatives (to within an average deviation of 2.2%). It does not follow the higher resonances in its present configuration, but there is no theoretical reason why it should not. The value found for the first formant corresponds to a low level component present in the filtered waveform.

Table III: Formant tracking results for fricatives.

Phoneme	Literature F1/F2	System Output F1/F2
/f/	1500/3000	570/1500
/θ/	2200/3500	575/2200
/s/	4000*/8000	600/2500
/ʃ/	2500/5000	500/2700
/h/	1800/2800	500/2800
/v/	1500/3000	500/1500
/ð/	2000/3500	535/1800
/z/	4000*/8000	545/2400
/ʒ/	2500/5000	545/2500

*Secondary resonance at 2500 Hz.

Table IV: Formant tracking results for nasals.

Phoneme	Literature Value F1/F2	System Output
/m/	330/2600*	470/2400
/n/	260/2300	450/2100
/ŋ/	260/2200**	450/2300

* Weak resonance at 800 Hz.
** Weak resonance at 600 Hz.

Nasals. The nasal consonants (/m/, /n/, /ŋ/) are produced by forming an oral closure and opening the velum. The closed oral cavity acts like a side branch resonator with the major radiation coming from the nostrils. This articulation causes an extra pole and zero to be introduced in the region of 1 KHz in the vocal tract transfer function. Nasal consonants are typically characterized by somewhat broader and more highly damped resonances than the vowels. The high frequency losses are attributable to the large surface are of the nasal tract that produces larger viscous and heat conduction losses (13).

Once again there has been no success in finding tabulated results for the formant frequencies of the nasals in children's voices. However, Table IV was constructed by comparing the data of Tarnoczy to the spectrogram sections of the particular speech samples being processed (14).

The second formant value output by the system agrees with the spectrographic section data and literature values to within 9%. It is believed that this is good accuracy when one considers the past performance achieved in formant tracking the nasals. It was difficult to decide where the actual peak of the spectral envelope was located on the frequency axis due to the previously mentioned problems of high pitch or glottal rate.

Once again, the baseline noise problem of the spectrum analyzer caused difficulty in judging the location of the first formant for the nasals. This problem is particularly severe since the first resonance is so low (less than 500 Hz). The values obtained from the formant tracking system are reasonable for /m/ and /ŋ/ due to the presence of another resonance in the F1 frequency range. /m/ has a resonance at 800 Hz as well as the predominant resonance at 330 Hz. This would cause the value reported by the system to be biased a bit higher than the actual value. There does not appear to be any way of filtering this extra resonance out unless special purpose hardware is added just for processing these nasals. A similar phenomenon is observed in /ŋ/ where the second resonance is at 600 Hz.

The results for formant tracking the second resonance were good, within 9% of the actual value. The results for F1 are inconclusive as of now, but the values obtained from the system do not appear to be unreasonable.

This concludes the test procedure as applied to isolated phonemes. As expected, the best performance was achieved when processing the vowels. This was due primarily to the fact that the filtering of the formant tracking system was tailored to the well defined formant structure of the vowels. The fricatives posed the problem of insufficient energy in the normal F1 range, but the F2 system output did follow the first resonance of the speech waveforms. Similarly, the nasals exhibited good second formant structure so that the system response was accurate. The results for the first formant are inconclusive when processing the nasals.

Analysis of Running Speech. In addition to the steady state sounds for which the results obtained have been already discussed, analysis was performed on running speech. Each subject was asked to enunciate several short sentences (e.g., we were away a year ago) to test the ability of the device to follow formant movements in each child's speech. These sentences were subjected to a complete analysis. Formant frequencies for F1 and F2 were calculated every 1/30 second. Spectrograms of these speech samples were then made and the system output was mapped onto the appropriate spectrogram. We found very good agreement between the two methods for the second formant. For the first formant the problem of determining accurately the formant frequency from a spectrogram has been discussed earlier. Consequently, it is not possible to say more than that good qualitative agreement was obtained. Since the system does a good job with steady state sounds in extracting F1, it is felt that the present analysis is adequate.

Conclusions

It has been shown that the formant tracking system in its present state of development accurately follows the time variations of the first and second formants in the speech of children for a large number of speech sounds. The fact that this system can extract these important speech parameters and display them in real-time implies that it can serve as the basis for an articulation training aid for the deaf. It would be presumptious to state that the current system is a useful aid that could take part in the every day speech training programs at a school for the deaf. The whole purpose of this study has been to show that the proposed formant tracking algorithm has the capability of following the formant transitions in the high pitched voices of children.

In order for a device to be a useful training aid it must be flexible, fun to use for both the student and teacher, easy to operate, dependable and cost-effective. It would be desirable for the training system to incorporate features that would attack all aspects of the speech-training problem. It would be easy to envision a computer-based system that would have a broad repertoire of teaching aids that could be selectively applied to the various teaching situations as they arise. There should also be a variety of practice and drill aids that the student can use on his own without a teacher's supervision. For this reason, the system should be so easy to use that it is in no way a hindrance to the teacher. In fact, it should be designed so that it is a desirable tool to use for both the teacher and student.

The areas of the speech training problem that appear to be most adaptable to computer system assistance are rhythm and timing, articulation of vowels (current system serves as the basis for this portion), production of fricatives, control of velum for nasalization of sounds, and finally, pitch control. The current system is directed at the articulation problem. Additional hardware would be necessary to monitor the other areas, and other types of displays might also be useful. Examples are: accelerometer mounted on the pupil's throat to measure the pitch of voiced sounds; vibro-tactile display for pitch feedback to the pupil.

Obviously the current system is a long way from the desired goal of a practical visual aid for the deaf. However, it must be noted that it has succeeded in providing a straight forward method of formant extraction with an accuracy that could not previously be obtained for the voices of children. The availability of the formant values in real-time makes an articulation training aid for children obtainable with some hardware additions to the existing system.

A subcommittee of the National Advisory Neurological Diseases and Stroke Council has stated, "there is a need to devise a system whereby each child in a deaf class can get much more individual drill in communication skills (particularly speech production) than is possible under present arrangements". This implies that a tutorial approach is suggested on an individual basis that is, the teacher would give his undivided attention to the problems and progress of the student. There is little doubt that the computer cannot replace this type of attention, but the aspect of much more individual drill could be handily monitored by a computer system. A computer-based system would be an excellent tool to provide the student with the necessary feedback during his hours of repetition and drill that is so important to learning effective speech skills. Hopefully the system can be programmed in a manner so that the drill and practice can be fun for the student. Games and cartoons that respond to speech parameters would be a good method to use, especially for preschoolers.

The basic advantages of a computer system are three-fold. First, even though it is relatively expensive, it can replace a myriad of special purpose devices, effectively reducing its apparent cost. Second, the general purpose nature of the computer makes it possible to think of a training system rather than a training device. It seems very unlikely that a single display will be found that is the solution to the speech training problem, so flexibility of a system approach is desirable since it can provide numerous aids that can be used as the occassion demands. Finally, the computer is not limited to speech training applications and it may serve many functions at the educational institution.

Whether or not the previous argument for a computer based speech training system is conclusive, one must agree that at least the computer system does provide a solution to the training problem. The engineer's job is now to design a flexible, dependable system that the student and teacher will want to use. The task is difficult at best, requiring knowledge of the psychology of the deaf, teaching methods, educational psychology, computer technology, and human engineering. A solution to these problems is not even proposed here, but a basis for one portion of a speech training system is set forth as worthy for consideration.

If the system is to function as an articulation training aid for the deaf, the displays must be improved in the many ways that have been discussed. The degree of flexibility required will make great demands on the computer's time (it is already pressed), therefore the auxiliary hardware will have to be expanded to further reduce the calculations the computer must perform in order to extract the formant values from the filtered speech signal. The ideal situation would be to do all the formant calculations in the speech processor and then interrupt the computer whenever a formant value is ready. With a few compromises a device to accomplish this goal can be constructed in a straight forward manner. In fact, with only a slight increase in complexity the device could be built as a self-contained unit that would provide an F1-F2 display. Such a device might be constructed for a relatively low price and could be made available to educational institutions as a special purpose training aid. Details of this proposed system have been worked out and will be presented in a future paper.

Acknowledgment

This work has been supported by Grant NS-08306 through the National Institute of Neurological Diseases and Stroke. This contribution is gratefully acknowledged.

References

1. Potter R.G., Kopp, G.A., Green, H.C., Visible Speech, Van Nostrand, New York, 1947.

2. Pickett, J.M., "Recent Research on Speech-Analyzing Aids for the Deaf"," IEEE Trans. on Audio and Electroacoustics, AU-16, 227-234 (1968).

3. Risberg, A., "A Critical Review of Work on Speech Analyzing Hearing Aids"," IEEE Trans. on Audio and Electroacoustics, AU-17, 290-297 (1969).

4. Thomas, I.B., and Snell, R.C., "Articulation Training Through Visual Speech Patterns," Volta Review, 72, 310-318 (1970).

5. Thomas, I.B., "Real-Time Visual Display of Speech Parameters," Proc. of Nat. Electronics Conf., 24, 382-387 (1968).

6. Levitt, H., "Speech Processing Aids for the Deaf: An Overlook," IEEE Trans. on Audio and Electroacoustics, AU-21, 269-273 (1973).

7. Niederjohn, R.J. and Thomas, I.B., "Phoneme Recognition of the Continuants in Converted English Speech," IEEE Trans. on Audio and Electroacoustics, in press.

8. Niederjohn, R.J. and Thomas, I.B., "Computer Recognition of Phonemic Segments in Corrected Speech," Proceedings of National Electronics Conf., 26, 83-88 (1970).

9. Niederjohn, R.J., Thomas, I.B., and Scott, D.E., "Use of the PDP-8/I for Formant Tracking and Classification of Speech Sounds," Proc. of DECUS Fall Symposium, 157-163 (1970).

10. Flanagan, J.L., Speech Analysis, Synthesis and Perception, Springer-Verlag, New York (1965).

11. P. Strevens, "Spectra of Fricative Noise in Human Speech," Language and Speech, Vol. 3 (1960).

12. Heinz, J.M., Stevens, K.N., "On the Properties of Voiceless Fricative Consonants," J. Acoust. Soc. Amer., Vol. 33, No. 5 (1961).

13. Fujimura, D., "Analysis of Nasal Consonants," J. Acoust. Soc. Amer., 34, 1865-1875 (1962).

14. Tarnoczy, T., "Resonance Data Concerning Nasals, Laterals and Trills," Word, 4, 71-77 (1948).

A MULTIPARAMETER VISUAL SPEECH DISPLAY*

by

Vassilios J. Georgiou and Ian B. Thomas**
Electrical and Computer Engineering Department
University of Massachusetts
Amherst, Massachusetts 01002

Summary. Channel data rate considerations suggest that a visual speech display may be useful for training the profoundly deaf to speak provided it is not too complex and provided there are no excessively rapid movements in it. A possible display device is described which presents simultaneously a number of speech parameters using position in the xy plane to convey articulation information (F1 vs F2), spot size to convey amplitude, and color to convey pitch information. To insure simplicity and reliability of operation, phase locked loops are used to extract the necessary speech parameters. These have been found to be excellent in pitch and formant extraction processes. For reasons of economy and ease of operation an unmodified domestic color television set is used as the display medium. The only connection to the television set is through the antenna terminals. Results obtained by using the device will be presented. The speech encoding device is economic, portable and very easy to use.

Introduction

Persons who are born deaf are not able to learn to speak as normal persons do because they have no feedback as to what sounds they produce when they use their vocal apparatus. If fortunate, they attend special schools for the deaf for 10 to 15 years and even then, after tedious instruction, they cannot produce high quality speech.

In a typical school for the deaf training takes place in teacher-student pairs. The teacher gives information on the desired position of the articulators for the production of a given sound. The student, on the basis of this information, makes utterances and the teacher judges the quality and accuracy of the produced sound and notifies the student accordingly. The feedback provided by the teacher in this fashion is not in real time and out of necessity it is not very sophisticated or accurate.

In order to provide accurate real time speech feedback, a number of devices have been constructed which analyze the speech sounds by extracting their basic parameters and displaying them in a visual or tactile display. A review of such devices was recently made by Levitt (1)

Despite their availability for some time, visual training aids for the deaf have not made any significant impact on the education of the deaf. This can be attributed to a number of shortcomings they have. They may be expensive, difficult to use, inaccurate or they may need to be adjusted to each particular speaker. In addition, some displays are too complex, overloading the trainee and some are too simple, thus giving inadequate feedback information.

The system described in this paper was designed in an attempt to overcome most of the problems encountered in the previous displays. The four basic speech parameters are extracted and displayed simultaneously in a visual display in a manner that is easy to perceive and follow.

System Description

The display used is basically an F1-F2 type display. The first formant frequency (by definition a formant frequency is a resonance of the vocal tract) controls the location of a spot along the x-axis of the display. The second formant frequency controls its position along the y-axis. Such a display is simple, yet informative. The first formant is related to the height of the tongue while the second formant is related to its front back position, thus giving a direct estimate of the position of this all important speech organ. In addition, encouraging results were reported by Thomas and Snell using such a display to train the deaf (2, 3).

Amplitude is displayed here as the width of the spot along the horizontal. Pitch (pitch is the frequency of excitation of the vocal tract, the frequency at which the vocal cords vibrate and it is also called the fundamental frequency or F0) is displayed as color.

In this fashion a multiparameter display is achieved in real time. Since position, size and color are used to convey information it is hoped that such a display will not overload the visual channel of the trainee. A block diagram of the display appears on Fig. 1.

The input speech passes through a preamplifier and then through prefilters suitable for each parameter to be extracted. These filters are followed by parameter extractions employing Phase Locked Loops (PLLs). Amplitude is extracted by fully rectifying the speech wave and lowpass filtering the result at 24 Hz. The resultant four speech parameters are converted into a video signal in the video formatter by suitable modulators driven by

---

* Also in the Department of Communication Studies.

**Please address all correspondence to Ian B. Thomas, Electrical and Computer Engineering Department, University of Massachusetts, Amherst, Massachusetts 01002 -- Telephone (413) 545-0769

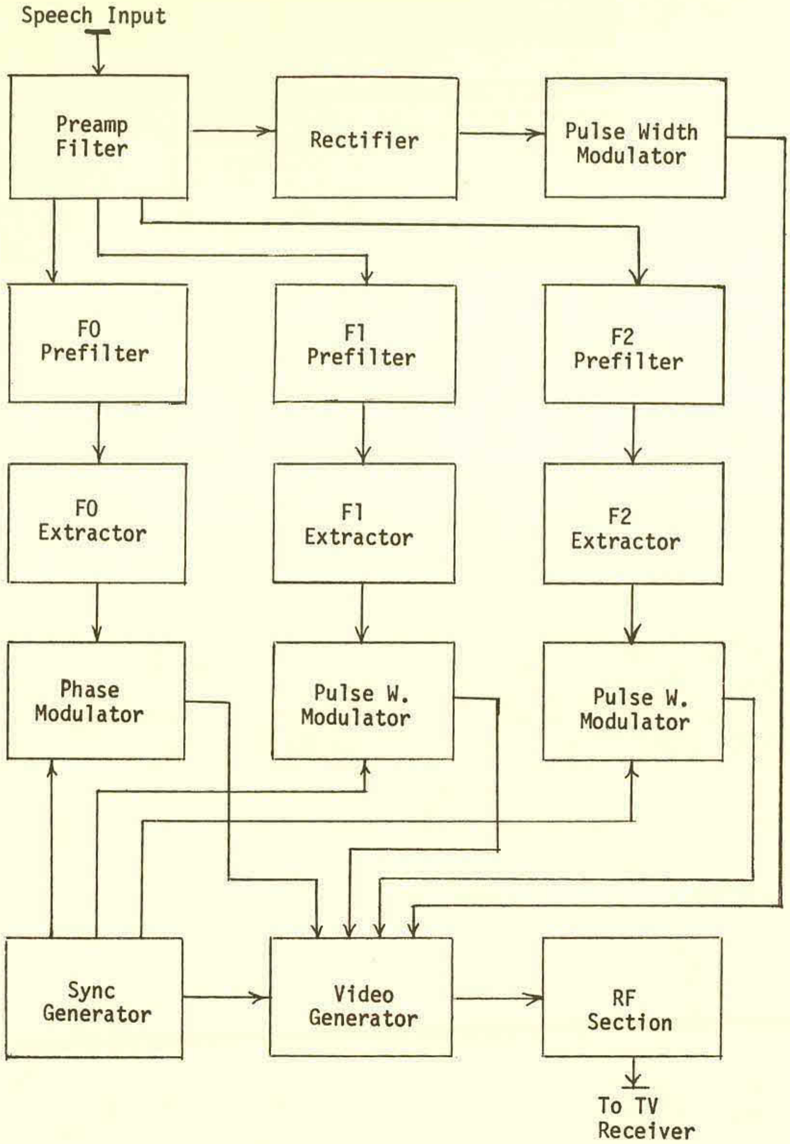


Figure 1 - System Block Diagram

the sync generator. The video signal then is modulated on the RF carrier which is transmitted with a transmission line to the antenna terminals of any unmodified commercial color television receiver. Such a receiver is widely available and it is the most inexpensive display medium with the necessary properties for this application.

Speech Parameter Extraction

A number of speech parameter extraction methods are available today but when they were evaluated for the purpose of driving the display generator, no one appeared to have the desired characteristics. As a result, a new method for formant tracking and pitch extracting was developed utilizing Phase Locked Loops (PLLs) which comes close to fulfilling the requirements of the display system.

The operation of the PLL during formant extraction is similar to its operation during FM demodulation. The spectrum of speech is a line spectrum because it is formed by the harmonics of the glottal frequency, the glottal wave being approximately a train of triangular pulses. The effect of the formants is to generate concentrations of energy peaks in the spectrum at frequencies corresponding to the formant frequencies. The movement of the formants then results in movement of the energy peaks in the spectrum which can be viewed as a kind of FM modulation.

An integrated circuit PLL was used and it was set to operate in first order with its center frequency at the middle of the formant range to be tracked. Before entering the PLL, speech is prefiltered in a spectrum tilt filter so that the effect of overlapping formant ranges will be minimized.

The output of the PLL containing the tracking voltage is lowpass filtered at 24 Hz and 36 dB/octave to remove components at high frequencies. The 24 Hz filter was chosen to optimally remove unwanted frequencies without introducing objectionable amounts of delay and averaging on the parameter trackers.

Pitch extraction is very similar to formant tracking although here the situation is exactly that of FM demodulation, making tracking easier and more accurate.

The formant extractors work well with a variety of input sounds, including whispered speech. They exhibit good accuracy, rejection of noise and adjacent formants and they quickly acquire and follow fast moving formants. Their output being a voltage level proportional to the tracking frequency is suitable to drive the video generator without further processing.

The Display Generator

A complete synchronization generation and RF transmitter-modulator is necessary in order to meet the constraint of using an unmodified commercial color TV receiver. In order to simplify design and to assure superior performance, an all-digital sync generator was designed using 20 TTL integrated circuits. It derives all timing information from a 3.58 MHz crystal oscillator which is also used to supply the color subcarrier. In the NTSC color television system, color information is conveyed as the phase shift of a 3.58 MHz subcarrier.

Pulse width modulations are used to define the vertical and horizontal coordinates of the spot. They are driven by the outputs of the F2 and F1 extractors respectively. They generate pulses which start at the end of the vertical or horizontal sync pulse depending on the modulator and whose length depends on the frequency the corresponding PLL is tracking. The outputs of the pulse width modulators are combined digitally and with a third pulse width modulator amplitude information is inserted in the combined signal. The pitch information in the form of the voltage output of the pitch extractor is used to drive a phase modulator which shifts the phase of the 3.58 MHz subcarrier by a maximum of 360° to produce all visible colors.

The color subcarrier and the pulse width signal are combined with the appropriate sync pulses and reference color burst forming the video signal. This signal is modulated on a low power 55.25 MHz transmitter for transmission on Channel 2 of the color TV receiver. A transmission line is used to avoid interfering radiation and connection is made to the VHF antenna terminals of the receiver. Due to the extensive use of integrated circuits and PC board techniques, the whole display system fits easily on a small table top cabinet and its overall cost is relatively low.

Discussion and Conclusions

The display system described in this report has been assembled and has performed well. The operation of the formant trackers is quite adequate when driving the display. There was a possibility that speaker to speaker variations, childrens voices, etc., might give marginal results in some cases. The testing done using sonagrams was limited to illustrative examples due to the tediousness of such tests. With a real time display a much more extensive testing is possible with no difficulty. Tape recorded voices of a child and of two women were tried. The performance of the trackers seemed to be just as good as with male voices. A tape recording of deaf speech was also tried as input. It had voices of various speakers with varying speaking capabilities. It was evident that the best sounding deaf speech gave a good coverage of the display area, indicating good control of formants. The poor sounding speech tended to cluster the area around the neutral vowels in agreement with previous observations. Tests were run with persons of normal hearing to examine how easy it is to produce sounds that would make the spot go to any location on the display. The results were very encouraging. The feedback provided by the display enabled the subjects to position the spot at any point on the display without much difficulty.

The feedback seemed to be quite suitable for consistent minimization of error. That is, once a sound was made that we in the desired direction, the subject kept on improving until the target position had been reached. Normal hearing subjects took only a few minutes to master the art of positioning the spot anywhere in the display area.

When tested with full words rather than vowels the display had a granular appearance due to the sampling inherent in the display (as was mentioned before, the spot is displayed once every 15 msec). In wide, slow excursions from the center point produced by some words this sampling was good since it gave timing information.

The amplitude display did not seem immediately useful. In most cases the fast moving spot was hard to analyze at first in real time. By repeating (using the tape recorder) the same word a number of times a kind of mental integration occured that tended to slow down the apparent spot movement. It seemed that one can learn to expect the major features of the display and the trajectory for a given word after a certain amount of repetition. Then attention could be diverted to the details such as size and color of the spot. This can happen within minutes and there are indications that once one becomes familiar with the display he will be much more capable of benefiting from it. It should be noted that using the display is very easy and one might say pleasant. There is a game-like quality in moving the spot about the screen, something like playing vocal tennis.

From these preliminary tests it is clear that the display system has a great potential to act as a display useful for training the deaf to speak. To be of any real value though, it should be integrated into a suitable training program. In the few attempts made so far to evaluate training aids for the deaf the results were encouraging but not spectacular. There are also a number of people who believe that, compared to conventional training, visual training aids offer no advantage. It is our opinion that the training method is of fundamental importance in achieving good results and that the effectiveness of an excellent feedback device might be compromised severely by the wrong training method. It will require an intelligent person with good knowledge of the speech field and a positive attitude towards the display to develop a good method of utilization. The value of suitable training can be judged from the cases of learning to type and learning the Morse code. In those areas there exist step by step methods which improve learning speed significantly.

A number of suggestions for improving the system can be made once it is seen in operation. First, one would like a facility for displaying each speech parameter alone for paying increased attention to any particular feature that might give difficulties. This can be easily done if the video signal is taken from the output of the horizontal pulse width modulator while the input to it is the desired parameter. Another potentially useful display would be the display of any parameter versus any other. To this end the spot display can be used after a simple change of the input to the vertical and horizontal pulse width modulators so that it is the parameters that are to be displayed. These are simple modifications of the existing equipment and should be added if they are thought to be of value later during the formal evaluation of the machine as a training aid. There are certain possibly desirable modifications that require extensive additions to the present equipment.

For example, in order to remove granularity, a digital memory can be used to store intermediate values of the point position and display them in order during the next frame so a continuous trace is generated. It is possible to design the memory so that a small number of bits will be required, about 5000, which can be provided by a small number of MOS memory chips at low cost.

An extension of this concept would be to have sufficient memory and appropriate control to store whole displays for later evaluation and review. One could conceivably store the target display and the student's two or three successive efforts to imitate it. Then the teacher would be able to point out features to be deleted or retained and discuss the display at his leisure.

Another feature that might be added is some PLL lock indicator which will blank the display in the case of no lock. As it is now, the display shows the trajectory of the loop as it acquires and looses lock. This might be of help but it also might be confusing.

Improvements in the type of the PLL used and in the loop filter will be certainly of value if they increase tracking capability. There is certainly room for improvement in this area. A not very successful attempt to improve loop performance was undertaken during this work and it is indicative of directions one might consider to increase loop performance. The VCO output of the PLL that was performing the formant tracking was divided by 4 and was fed into a PLL designed to operate at one quarter of the center frequency of the first PLL. It was hoped that the division process will provide filtering action by averaging out four consecutive cycles and thus give a better input signal to the second PLL. This scheme unfortunately did not give better results than just one PLL alone and it was abandoned.

In conclusion it can be said that the display described establishes the ground for low cost, versatile high performance visual speech training aids for the deaf. In addition, it is hoped that it will spur interest in utilizing the versatility and low cost of the commercial color television receiver as a display medium for a variety of purposes.

Acknowledgement

This work has been supported by a Public Health Service Grant NS-08306 through the National Institute of Neurological Diseases and Blindness.

References

1. Levitt, H., "Speech Processing Aids for the Deaf: An Overlook," IEEE Transactions on Audio and Electroacoustics, AU-21, 269-273 (1973).

2. Thomas, I.B., and Snell, R.C., "Articulation Training Through Visual Speech Patterns," Volta Review, 72, 310-318 (1970).

3. Thomas, I.B., "Real-Time Visual Display of Speech Parameters," Proc. of National Electronics Conf., 24, 382-387 (1968).

BANQUET ADDRESS

by

Captain G. E. Welch
Director, Access for the Disabled
Central Council for the Disabled
London, England

Captain Welch entered the British Army from Military Academy at the beginning of World War II. He spent seven years in the Far East, transferring to the Indian Army in 1943, and saw service in India, Assam, Burma, Malaya, French Indo China and China. He suffered a severe attack of 'polio' in 1947 while in Rangoon, and after three years in various hospitals was retired by the Indian Government in 1950.

After ten years in industry and commerce, he entered Government service and took an appointment in the Research Department of the Home Office. While at the Home Office, he read Law and Social Science at London University. In 1967 he transferred to Local Government service and took a senior administrative appointment with the County Borough of Northampton. He joined the Central Council in 1970.

There is one great advantage in being disabled--instead of having to stand up and give a speech, one can sit down in comfort and just talk. Seriously though, it is not only a pleasure, but a privilege to be speaking to you tonight. Having heard such excellent speeches at the conference sessions today, I trust you will not feel let down by my nontechnical speech tonight.

When Professor Jackson first invited me to be guest speaker at this banquet, I expressed doubts as to the suitability of my subject. I pointed out that there were many large and powerful organizations in America dealing with the physically handicapped, and they all had the expertise to solve the access problems of the disabled. I also stressed that America had its Code of Building Practice just as we in Britain have the British Standard Code of Practice. I also emphasized that I was certain the Americans were just as aware of the needs of the disabled as we were in Britain.

Professor Jackson then pointed out something I had overlooked, namely that the United States has Federal Government and State Government. This meant that in day-to-day affairs each state was practically autonomous, with the result that it was possible for one state to differ greatly from another. He made me feel that he would like to see many of the things we do at the Central Council being done in Kentucky.

Thus I was convinced, and felt I had three good reasons for coming: (1) by being an observer at the three-day conference I would obtain a tremendous amount of information on electronic aids for the disabled, to take back to Britain; (2) I would derive great pleasure if, as a result of my talk this evening, and the discussions I shall be having with Professor Jackson after the Conference ends, I can help to promote interest in the many and varied problems of the disabled in Kentucky; and (3) my wife and I could stay on after the Conference and avail ourselves of the kind offer of Professor Jackson and his charming wife to show us the Blue Grass State.

I have recently completed a tour of the main towns in Wales, lecturing on the access problems of the disabled and the design criteria necessary to eradicate or at least minimize these problems. I am not going to repeat that lecture as time will not allow me. One cannot give an hour-and-a-half lecture followed by a half-hour slide show, in the time allowed for a banquet speech. In any case, I am certain you are all aware of the architectural problems facing the disabled, and are in a position to obtain advice on necessary design criteria from your national organizations. I might add that should anyone wish to discuss the problems with me during my stay in Lexington, I would be only too pleased to do so.

I have decided that the best way to use the short time at my disposal, is to give a brief description of the work of the Central Council and show how it helps to ease the problems of the disabled, with special emphasis on access. The Central Council for the Disabled is a national action group for all physically disabled people. It acts as a coordinating body of voluntary organizations to bring pressure to bear on central government and local government, aiming to improve the living conditions and environment for all disabled people. It concentrates on Access, Housing, Holidays, and Legal and Parliamentary affairs. The Central Council maintains a Travelling Exhibition of Aids for the disabled, and publishes a monthly bulletin.

Nearly 400 organizations are affiliated to the Central Council, half of these being voluntary organizations, and half being statutory bodies. The voluntary organizations include such national bodies as Polio Fellowship, Spastics Society, Multiple Sclerosis Society, Muscular Dystrophy Society, etc., and all County Associations for the Physically Handicapped. The statutory bodies are made up of County Councils, County Borough Councils, and Borough Councils. The councils have elected members who are responsible for their bye-laws, but nothing else. All legislation emanates from the central government and the councils enforce it. This is possible in a small country like Britain.

These councils have a director of social services who, as his title implies, controls all social services which naturally includes welfare services. This department provides the services for the physically and mentally handicapped.

Just as you in America have good and bad states, we in Britain have good and bad local authorities; but I am glad to say that over the last two years, as a result of publicity campaigns carried out by the Central Council and its hundreds of affiliates, the number of backward authorities is diminishing. The turning point for the physically handicapped in Britain came in June 1970. Prior to this, the welfare departments of many local authorities accomplished a great deal for the disabled, and of course many didn't, because it was not mandatory. However, on June 1, 1970, the Chronically Sick & Disabled Persons' Act was passed in Parliament. It was a Private Members Bill which became law with the clear intention of Parliament, that it would be implemented in an atmosphere of co-operation. It contains 29 sections, all vitally important to the disabled person. I'll quote just one of them briefly: 'It is mandatory that any new building to which the public are to be admitted whether on payment or otherwise, shall be made accessible to the disabled." The mandatory powers of the Act did not include existing buildings, much as both sides of the government would have liked, as the cost would have been astronomical. However, there was expectation on both sides of the upper and lower Houses of Parliament, that existing buildings would be adapted wherever possible in the spirit of the Act.

In this matter, the Central Council, as a pressure group, has been able to persuade many a local government authority and private business concern to carry out adaptations, thereby enabling the disabled person to lead as normal a life as possible. One has to be practical and realize that some buildings by virtue of age and architecture, cannot be adapted, even with all the good will in the world.

The passing of the Act is only the first step--it doesn't wave a magic wand and produce men, money and materials overnight. But at least central government and local government have become aware of the needs of the disabled, and that is a step in the right direction.

Access is the operative word running through a disabled person's life. In fact it governs his way of life, because in whatever he tries to do, invariably "access" crops up. If he wishes to take a holiday he must have access to a hotel, motel or guest house. If he hasn't his own private transport he needs access to railway stations, bus stations and airports, as well as to the means of transport. He cannot get a job if the place of employment is inaccessible. If he is denied access to the buildings used by the public, whether they be places of entertainment, trade and commerce, education, or recreation, and if he is prevented access to a town's facilities by virtue of newly planned pedestrianisation schemes, then he cannot enter fully into community life. He is thus prevented from leading as normal a life as possible, which is morally wrong.

Even the question of living accommodation involves access. There are many people immobolised as they haven't access from one part of the house to another.

When looked at in this light "access" assumes gigantic proportions and is seen to be the word that really governs the disabled person's life. To many people, an access problem is just some architectural obstacle such as a door too narrow for a wheelchair, or some steps that need a ramp, and once the necessary adaptations are carried out, the problem is solved and can be forgotten. This is not so. The solving of access problems, and ensuring that the disabled lead as full a life as possible calls for much more. National organizations such as the Central Council have to be permanent pressure groups; keeping central government and local government aware of the needs of the disabled; co-ordinating the efforts of voluntary organizations so that they can act from strength; making everyone (including the general public) aware of what is being done and what can be done for the disabled; keeping up a good publicity and information service; and maintaining a close liaison with central government ministers and local government directors of social services.

I think the easiest way to demonstrate the Central Council's actions is to briefly describe various projects and developments. Whenever I am asked to speak on access problems and design criteria by any local authority, the director of social services always arranges for a well-balanced audience; one half being made up of professional workers of the local authority, e.g. the director and his staff, architects, engineers, surveyors, social workers, and occupational therapists; the other half comprising members of voluntary organizations concerned with the disabled.

The local authority representatives are the ones who can authorize and carry out the necessary work to make buildings and places accessible to the disabled. The voluntary organization representatives know the needs of the disabled and can help their cause by having a good liaison with the local authority. Also, on the odd occasions when some authorities lag behind in carrying out their duties, they can be vital pressure groups to ensure things get done. By making these two groups of workers aware of the access problems and the design criteria necessary to solve them, whether concerning universities, colleges, residential schools, theatres, sports centers, swimming pools, public buildings, places of employment, or pedestrianisation schemes, and also informing them of the architectural publications and government circulars available giving design particulars, the disabled can be helped.

The Central Council employs a consultant architect who specializes in designing and building for the disabled. Even in his private practice he and his partners do nothing else. His advice is given free to everyone as we pay him a retaining fee. He has given advice and help to many local authority architects and private architects, either by way of written letter, or the vetting or amending of plans and drawings. The following is an example of the service given:

(1) Design of layout for pedestrianisation schemes so that disabled drivers have access to areas where normal drivers are not allowed.

(2) Purpose-built public conveniences.
(3) Building of sports centers and swimming pools.
(4) New shopping redevelopment schemes.
(5) New integrated schools.
(6) New town developments--the needs of the disabled.
(7) Adventure playgrounds for the disabled.
(8) Access problems at places of education.
(9) Day centers for the disabled.

Some of the enquiries have entailed our architect and myself visiting the sites for discussion on the various projects. The advertising of the service has resulted in many access problems being solved all over the country.

Many organisations throughout the country, both statutory and voluntary, have availed themselves of our comprehensive guide-making instructions. To date nearly 100 guides have been produced on towns as far apart as Aberdeen in Scotland and Plymouth in Southwest England. At present another 40 are under construction. There is no need for me to emphasise how these guides help a disabled person with his access problems.

We have just completed a Guide to Public Conveniences in England and Wales. We wrote to 700 local authorities for particulars of facilities available. The data supplied was sufficient for a guide.

A short while ago we completed a Guide on every University and Polytechnic College in the country. This shows what faculty buildings are accessible to wheelchair and ambulant disabled.

We have just completed a Guide to the London Underground System. This is invaluable to the disabled as the facilities of every station are shown.

Last month we wrote to more than 200 railway station managers asking them to fill in a questionnaire on facilities available. The replies received are encouraging, and the data are sufficient to produce a Guide to British Rail in the near future.

We publish a book each year entitled "Holidays for the Physically Handicapped." This gives particulars of hotels, guest houses and holiday camps able to cater for the disabled. It is kept up to date by one of our staff, a physiotherapist, driving around the country for most of the year vetting such establishments. The book also contains similar information on 15 European countries.

Each year during the first week in October we organise a "Help the Disabled Week", and the subject last year was "Access for the Disabled." Information kits were sent to our 400 affiliates. These kits contained information on the international access symbol, the minimum criteria necessary to display the symbol, and instructions on how to give the symbol as wide publicity as possible. These instructions included a specimen draft of a letter to the press, and the way to organise surveys of all types of public buildings. The press and publicity media were informed of these surveys, and the co-operation of the editors of the particular local papers was obtained. This meant that our affiliates had a blitz on 400 towns. All buildings which came up to the required standard were given an access symbol to display. The publicity arising from this week snowballed, and during the following months I received dozens of letters from statutory bodies and private enterprise, asking for particulars of what was necessary to make their buildings accessible.

The topic this year is to be "Employ the Disabled" and the same sort of publicity blitz will take place.

Besides maintaining a close liaison with local government directors of social services, we have a good liaison with central government ministries. There are three ministries we deal with regularly: Department of the Environment, Department of Employment and Department of Health & Social Security. We have a good rapport with the top permanent civil servants of these ministries, and when the occasion arises the minister himself never refuses to meet us.

Naturally a government cannot deal with individuals --it hasn't the time--but it will deal with a national organisation such as the Central Council.

All information we receive we pass on in our monthly bulletin. The bulletin, which averages twelve to sixteen pages, contains interesting information pertaining to the disabled, taken from local papers all over the country, information received direct from the various government departments, and legal and Parliamentary news supplied by our own Legal & Parliamentary Committee. The bulletin has a publication of 1,100 copies. This may not sound like much, but 400 of these copies go to our affiliates who in turn give the information to thousands of disabled people. Another 400 go to small groups of disabled people. The editors of 60 magazines on disability also receive copies, and naturally they spread the news far and wide.

Our Legal & Parliamentary Committee consists of specialists in certain aspects of disablement. The Committee worked very closely with the Member of Parliament whose Private Members Bill became the Chronically Sick & Disabled Persons' Act. During the period of preparation an ad hoc committee was set up consisting of the Legal & Parliamentary Committee, one or two interested specialists outside of the Committee, and Members of the House of Commons and the House of Lords. For a period this ad hoc committee met weekly in the House of Commons on the day prior to the committee stage of the Bill, and when the committee stage was completed, at intervals prior to the report stage and the third reading in the House of Commons and its passage through the House of Lords. The secretary of the Legal & Parliamentary Committee attends Parliament on any day the programme contains anything affecting the disabled. She receives a copy every day of "Hansard", which is a verbatim report of everything said in Parliament. All reports of interest to the disabled are extracted and published in our bulletin. Thus the disabled are fully aware of what is being done on their behalf. There is an All Party Committee on Disablement in the House of Commons. This is non-political, comprising members of government and the opposition in equal numbers. This is ideal, as under conditions such as this, politicians become human. Our Legal & Parliamentary Committee meets with this committee a number of times a year which enables it to put forward any points it feels would help the disabled.

To show the close liaison we have with the central government, I give the following example: Some months ago one of the members of the Parliamentary All Party Committee read an article in the press criticising certain cinemas in their attitude to wheelchair patrons. He contacted the Cinematograph Exhibitors' Association, and suggested a meeting. This was agreed to, as they were averse to public criticism. My Director and I met senior executives of CEA and two members of the Parliamentary Committee in the House of Commons. This led to further meetings between the Central Council and the CEA. The final result is that the CEA will be sending us a draft Code of Practice for vetting. Once we agree to it, the Code will go to the General Committee of the CEA for ratification. Once ratified, the Code of Practice will be published in the CEA's newsletter to managers of every cinema in the country. This will stabilise matters, and ease the access problems of the disabled.

The same sort of liaison occurs with local government. A short while ago my Director and I met the planning committee of the Association of Municipal Corporations. This Association is the executive body of all town councils. We discussed the ogre of pedestrianisation, emphasising that although pedestrianisation schemes help to improve the environment of city centres, the needs of severely disabled people must be taken into consideration if they are to take part in the complete life of the community. Many disabled persons, because of lack of mobility, are dependent upon private transport whether as a disabled driver or disabled passenger. We stated our requirements and pointed out that it is possible with sound planning to allow disabled drivers and passengers access to precincts that exclude other traffic. This was accepted, and resulted in a circular being sent to town councils throughout the country, asking them to take heed of the needs of the disabled when planning future schemes.

Dealing with access problems and trying to ensure that the disabled lead a reasonably normal life is an ongoing thing. There has to be continuous liaison with central and local government, regular publicity of what is being done and what should be done, a good information service to educate organisations, both government and private, on design criteria and what is available for the disabled, also to let the individual disabled person know what he is entitled to, what help is available, where he goes to get help or equipment, and what legislation has been passed to help him. Through our liaison with, and lobbying, of government departments, various pieces of legislation have been passed, and resulted in our being able to publish circulars informing the disabled of their entitlement.

To keep up the pressure in publicity, and ensure that things are being done, we have regular contact with our affiliate organisations. Sometimes we feel there is something that needs doing in their area, and they with their local knowledge are equipped to do it--so we encourage them to take the necessary action. On other occasions the affiliates feel they need our help--they may wish us to approach their local authority on a certain subject, or they may think a problem is too big for local authority to deal with, and one of the central government departments should be contacted. This two-way action helps to keep things moving.

Earlier I mentioned our Travelling Aids Exhibition. While this is not basically access, the aids exhibited are aids to living, and make a disabled person more mobile, and I feel there is a close affinity between access and mobility. This exhibition is on the road from April 1st to October 1st. Any local authority or voluntary organisation may book it for one week free of charge. The organisation booking it is responsible for providing a suitable hall, for arranging publicity, and for sending out invitations to local doctors, nurses, social workers, voluntary organisations, hospital staff, and disabled people and their families. Our chief aim is toward a professional audience, as one occupational therapist or social worker can of course pass on the information to many disabled people. It is essential that these professionals are made aware of the aids and equipment available to increase the mobility of the disabled. This exhibition is informal, visitors wandering around examining the aids and discussing points with our occupational therapist in charge. Firms are only too pleased to lend items for exhibition as it is good free publicity.

I mentioned earlier on that even in housing, access problems occur. Our Housing Department gets many applications for purpose-built housing or adaptations to their existing property, to enable them to lead a life of less stress and strain. Sometimes we can get local authorities to help, but only in a minority of cases, as most authorities have housing waiting lists. We have a good rapport with housing societies and housing associations all over the country--there are hundreds of them. They are non-profit organisations supplying flats, houses and bungalows on a cost rent or co-ownership basis. When they build new properties they build a certain percentage suitable for the disabled. When they take over old, large townhouses and convert them into flats, they endeavour to let us have ground floor flats for the disabled.

Under the Chronically Sick & Disabled Persons' Act it is mandatory for a local authority to carry out adaptations to a property occupied by a disabled person. This applies to owner occupied property, as well as council owned property.

I realise that the government set-up in America is totally different to that in Britain. Your federal and state system is going to call for a different approach. In any case, your country is so large it would not be possible for an authority in Washington to have close liaison with small organisations far and wide, as we have in Britain. However, that doesn't mean each individual state cannot organise services for the disabled. You have many large and powerful voluntary organisations for the disabled. These organisations have the knowledge and expertise on design criteria for the disabled. I am certain it would be possible for such organisations existing in Kentucky to develop close liaison with the departments of environment and employment in the state government, just as we have done with our central government in London. I appreciate you have no legislation at the moment comparable with our Chronically Sick & Disabled Persons' Act, but coordination of the powerful voluntary organisations, good publicity, and lobbying of the state government should have tremendous results.

You in Kentucky are a member state of a powerful country that puts men on the moon with the essence of ease. Anyone with that amount of drive and initiative should take the problems of access for the disabled in their stride.

APPLICATION OF ELECTRONIC PROSTHESES IN REHABILITATION AND EDUCATION OF MULTIPLY HANDICAPPED CHILDREN

by

F. A. Harris, F. A. Spelman, J. W. Hymer, and C. E. Chase
Department of Physiology and Biophysics/Regional Primate Research Center
University of Washington Medical School
Seattle, Washington
and
Program for the Physically Handicapped
Seattle Public Schools

A fundamental distinction should be recognized between electronic prostheses which are intended to substitute for lost or impaired exteroceptive sensory function (e.g., laser cane, ultrasonic "glasses", Opticon) and those which are used as assistive or training devices to promote rehabilitation of individuals with central nervous system dysfunction. Dealing with the postural and mobility problems of cerebral palsied children requires a bioengineering program directed towards development of the latter, which we feel presents a significant challenge to applied neurophysiological theory and electronics technology. We are currently working with a public school district Special Education facility, as a major portion of the Program for the Physically Handicapped, attempting to develop training devices and accessories to meet the full spectrum of rehabilitation and educational needs of cerebral palsied children.

By the end of the second year of our program (1972-73) we will have finished training devices or systems for: 1) teaching head control and stability; 2) teaching limb position control; 3) prespeech breath and speech articulation training; 4) visual perceptual and sound-symbol association training; 5) gait pattern and foot alignment training; 6) positioning arm and head to enable self-feeding; 7) teaching expressive language through printed communication; and 8) preventing drooling and training tongue movements for speech. All of these devices emphasize improvement of sensory feedback as their primary functional characteristic (1). To date, extensive data has been collected to document the effectiveness of the first four devices; the others were finished too late in the school year to permit full evaluation and/or insufficient school personnel were available to use them in therapy on a routine basis. In the following sections, each device or system will be described, the way in which it was or will be employed in therapeutic activities will be discussed, and the data analyzed at this time will be presented to permit evaluation of the benefits accruing to users. (At the Conference, slides were shown of each of these devices or systems and movies were shown of children engaged in therapeutic activities utilizing them.)

1) Head Control System (HCS)

This system is essentially the same as described in the Proceedings of the 1972 Conference, with two additions (2): 1) A DC oscilloscope has been added to the system for visual feedback of head position. Vertical deflection of the beam corresponds to front-back movements of the head, and horizontal deflection to side-side movements of the head; in treatment, the child attempts to keep the beam within a circular target at the center of the screen. 2) A three-channel MFE stripchart recorder has been used to provide permanent records of head movement by monitoring the transducer outputs. The latter permits far more objective determination of the efficacy of the Head Control System than did cinematography, although filming is still carried out at regular intervals so that educators can view the child as well as stripchart records, graphs, and numbers representing his performance.

Ten cerebral palsied children, ranging from 7 to 16 years old, were treated with the Head Control System during the 1972-73 school year. The treatments were one-half hour in length, and the number of treatments per child ranged from 19 to 77. The children practiced therapeutic activities (bringing the head back to the neutral position after a deliberate offset by the therapist; moving the head away from the neutral position themselves and returning to it; and, attempting to hold the head at the neutral position as long as they could) for the first 25 minutes of each session, and then the last 5 minutes of activity were recorded for data collection. In order to vary the task and prevent boredom in the early training stages, the children sometimes worked with auditory feedback and at other times with visual feedback. When a given child had developed the ability to stabilize the head for at least 20 seconds, he was placed on the movie projector reinforcement system for subsequent training. Here, he was motivated to stabilize his head by making his viewing an interesting movie contingent upon his keeping his head within adjustable limits of tilt. At first, the child controlled the projector with a single channel of movement; as he gained proficiency, both channels were used simultaneously to control the projector actuation. Provision was made for progressively shaping finer and finer control of head movements in that as the child's skills improved through training, the tilt tolerance for the projector actuator was decreased by 5 degrees (2.5 degrees in each direction) for every change in sensitivity setting.

Results (HCS)

Data was analyzed in three different forms: 1) "eyeball" comparison of stripchart recordings was made at various times during the treatment period; 2) "duration histograms" were constructed to depict the relative number of occurrences of intervals of stability of various length during a particular treatment period and these were compared on a weekly basis to assess short-term improvement; and 3) bar graphs were constructed to depict the percentage of the total recording time during which the head was held stable, on a daily basis, for the entire period of treatment. Figures 1-3 provide samples of each of these.

In Figure 1, the baseline record (A) shows incessant, high amplitude (90 degrees) and rapid (4 cycles per second) movements of the head, both front-to-back and side-to-side. Control is slightly better in the FB direction than in the SS; the

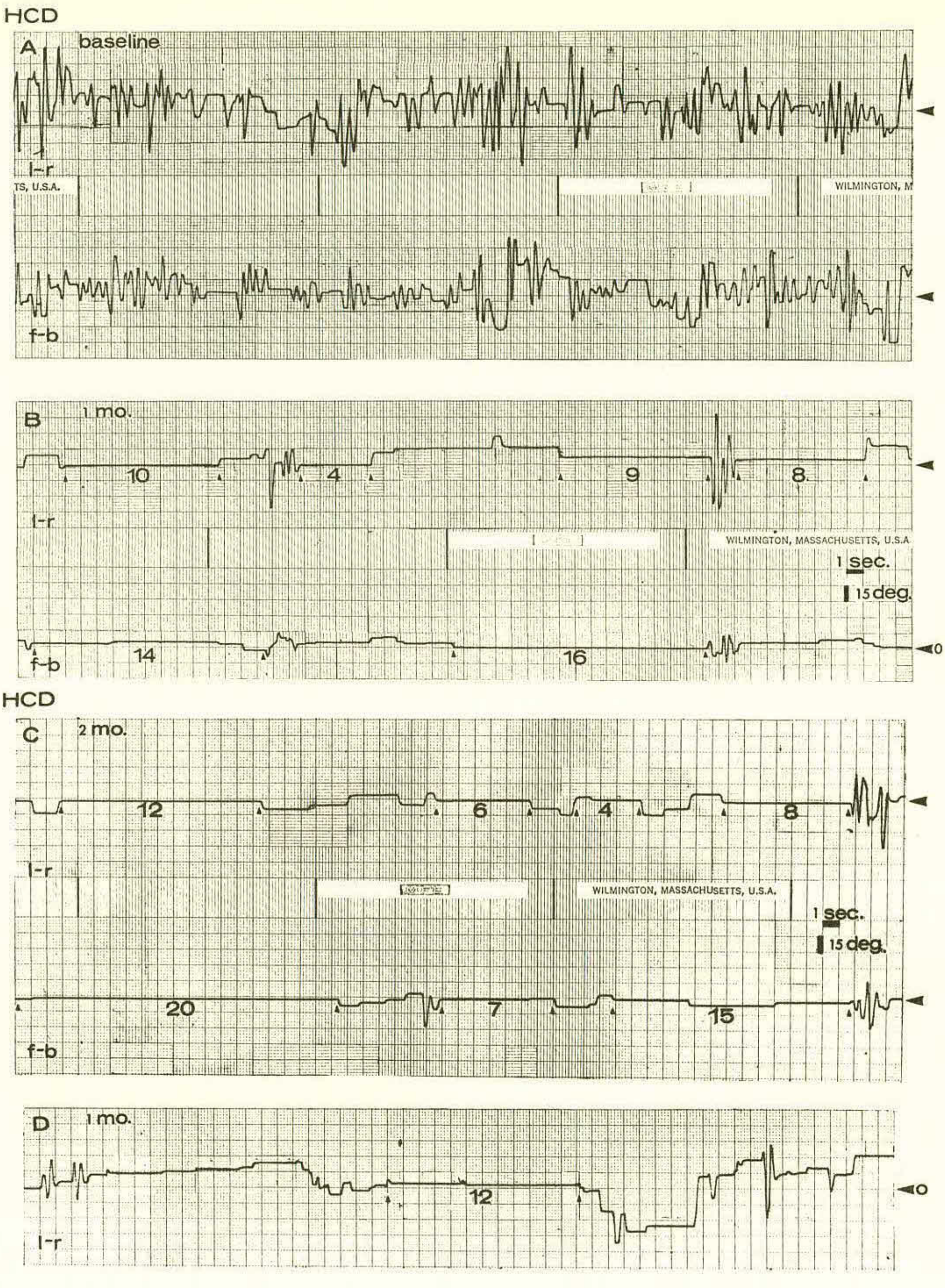
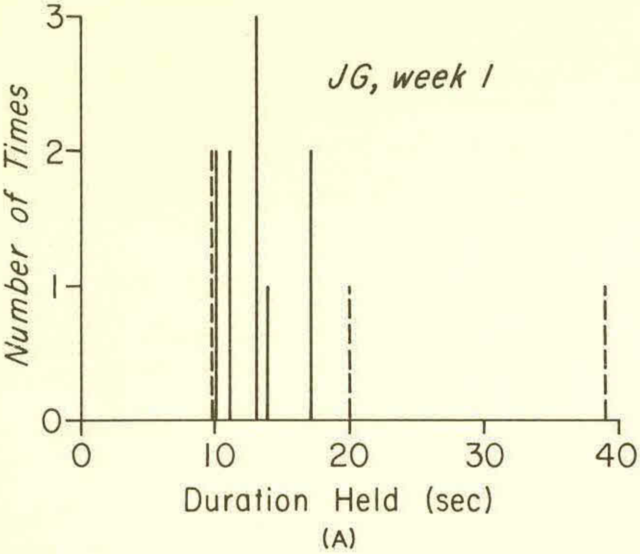
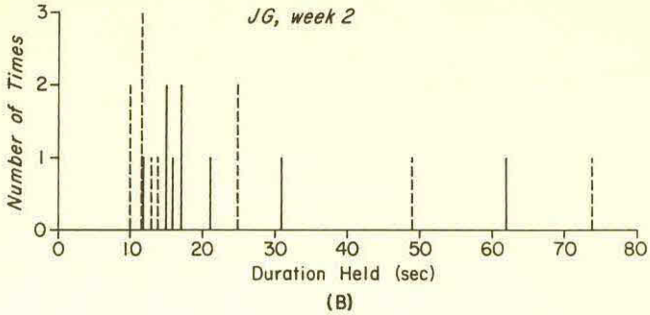


FIGURE 1A-D. Stripchart recordings for head control system.

longest period of simultaneous stability in both channels is one second. After one month of treatment (B), the improvement is striking. There are extended stretches of stability in each channel, and two periods of simultaneous stability approximately 10 seconds long. In a 55 second period, the head makes large-scale movements only twice. After a second month of treatment (C) the child is able to hold his head within 5 degrees to either side of the neutral position in both channels for the entire period shown, except for a few seconds. The 20 second stretch of perfect stability at the neutral position in the FB channel is particularly striking.

Record D is for another child with far more severe movement problems, and represents movement in the SS plane only. Initially, the child could not hold his head still for even one second. His head was held upright by a restraint attached to his





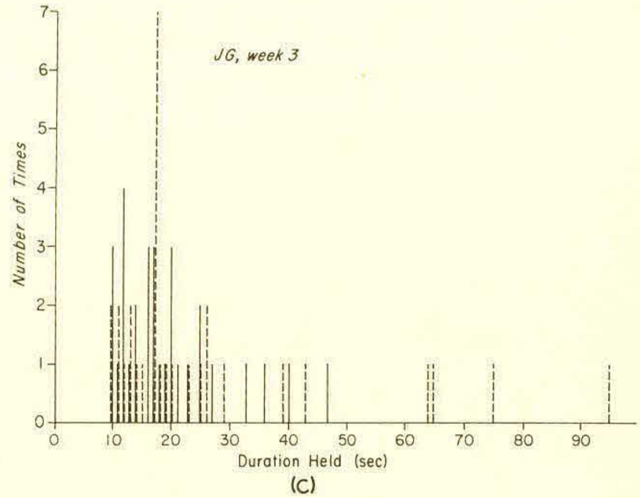


FIGURE 2A-C. Duration histograms summarizing first three weeks' progress for child using head control system.

chair throughout the day, and our goal was to teach him sufficient control to free him from dependence upon this support. It was necessary to work with each plane of tilt separately in this case. The child had virtually no head control without the device; with it he had learned to hold his head steady at the neutral position for a 12 second interval after one month's treatment.

Figure 2A-C consists of "duration histograms" representing the first three weeks of treatment for one of the children. The graphs show the number of times which the child was able to hold his head steady for intervals of various lengths; at this stage of treatment, the FB and SS channels are considered separately. Each graph represents fifteen minutes of recording (five minutes at the end of each of three days of treatment per week). There is a clear progression towards longer periods of stability from week to week; control of the FB channel is better than that of the SS channel, which is typical of all of the youngsters involved in this program.

Figure 2D shows the percentage of recording time during which the child is able to first keep her head within 20 degrees of neutral and then to keep the projector running, with both channels simultaneously governing the projector controller, on a daily basis for several months of treatment. The arrowheads at the top of the graph indicate points at which a change in the procedure took place; the first change was the introduction of the movie projector reinforcement system, and subsequently changes were made in the tilt tolerance setting on the projector controller. Learning is taking place in each phase of the treatment. There is an initial increase to 83% stability time with auditory feedback only, then a drop as the task is changed with the addition of the projector and eventual recovery to higher levels of performance, then subsequent drops and recoveries each time the sensitivity setting is changed. At the end of the period under consideration, the child is able to keep the projector running for 75-80% of the time with a tilt tolerance of 15 degrees from the neutral position in both directions.

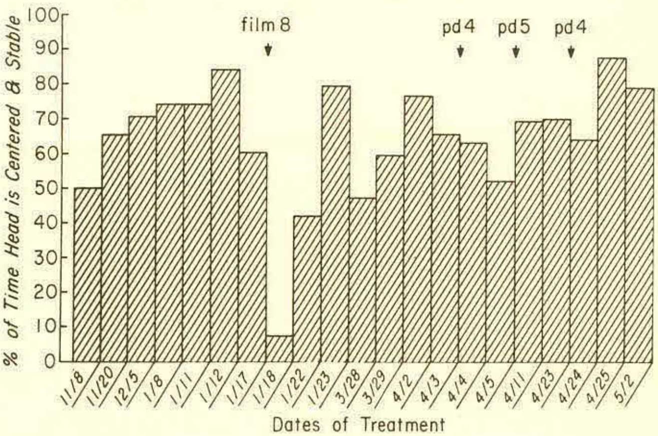


FIGURE 2D. Summary graph showing percentage of recording time during which child maintains proper and stable head position.

2) Limb Position Monitor (LPM)

The Limb Position Monitor and the treatment procedure utilizing it are also the same as were reported at the 1972 Conference. As an addition, stripchart recordings of movements at the elbow were made during the current school year to document the effectiveness of this device; this was accomplished by monitoring the transducer output. Also, a new "tracking" version of the Limb Position Monitor (designated LPM-T) was developed and put into use for an advanced level of limb control training. The LPM-T employs two servomotors driving concentric shafts, one actuated by a function generator and the other by input from the joint position transducer. The function generator is used to move a "target" over various ranges of motion, at various speeds, and in different modes (sine wave, square wave, etc.). The child moves his arm in such a manner as to keep the pursuit pointer "on target" while the target moves in a predictable fashion. After having been trained to hold his elbow joint immobile for at least 10 seconds, and then practicing moving the forearm repetitively over various ranges while using the original version of the LPM, the child graduates to the LPM-T and from then on works in a more dynamic situation. (From time to time, he returns to use of the LPM to determine if the dynamic training has further improved his static control.)

When the LPM-T is used, the output of the function generator is displayed on one channel of the stripchart recorder and the movement of the joint is displayed on the adjacent channel for comparison.

Six cerebral palsied children, ranging in age from 10 to 18, worked with the LPM-T during the 1972-73 school year. The total number of treatments, again one-half hour's length, ranged from 26 to 39. They practiced tracking the target moved at various speeds and over various ranges for the first 25 minutes of each treatment period, and then attempted to pursue the target moved over a 135 degree range at 3 and 6 cycles per minute (sine wave) and 3 cycles per minute (square wave) while recording was carried out.

Results (LPM)

Data was analyzed in two different forms: 1) "eyeball" comparison of stripchart recordings was made at various times during the treatment period; 2) the percentage of "good quality" movements was determined and compared, for each of the modes of movement, for the initial, middle, and final period

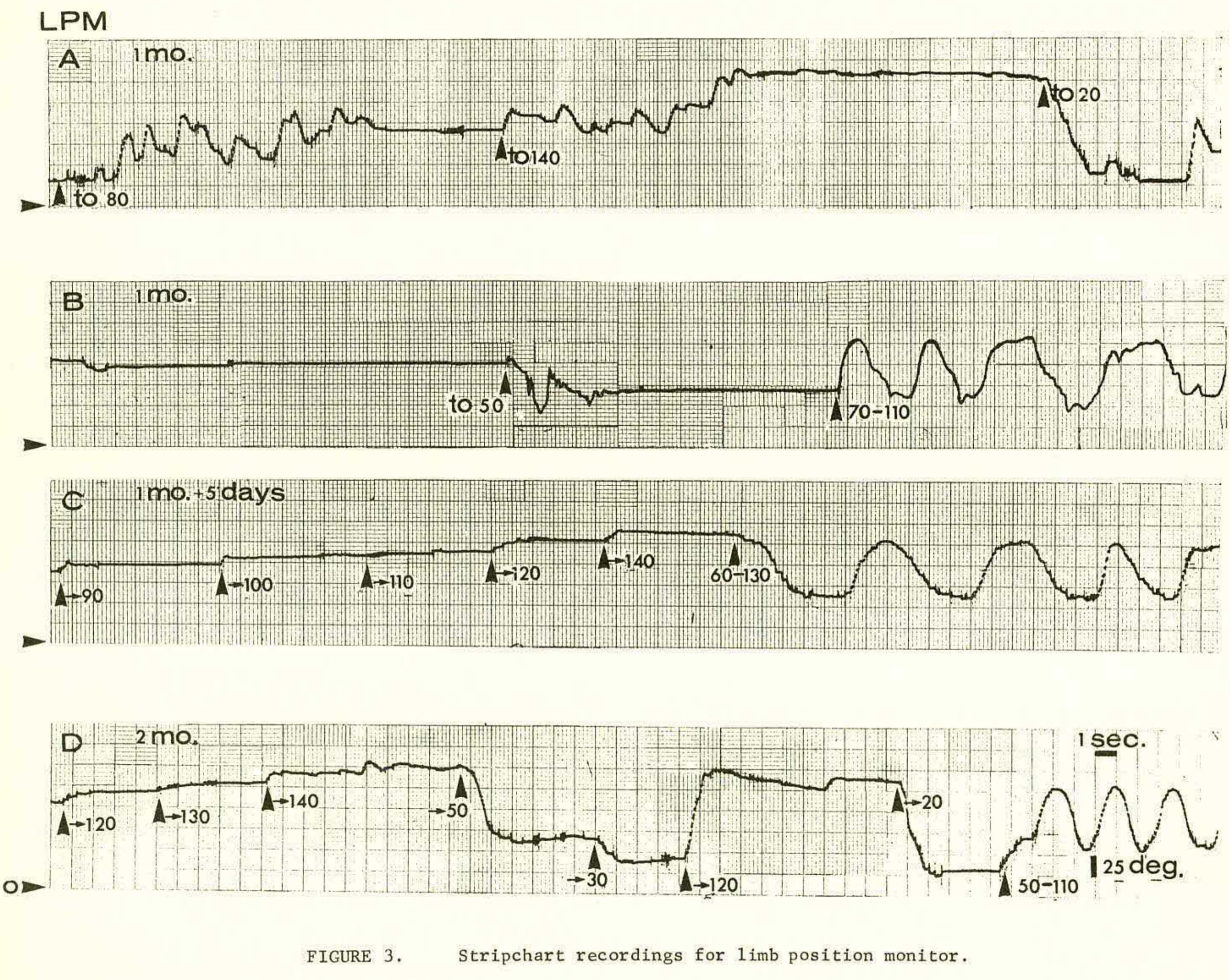


FIGURE 3. Stripchart recordings for limb position monitor.

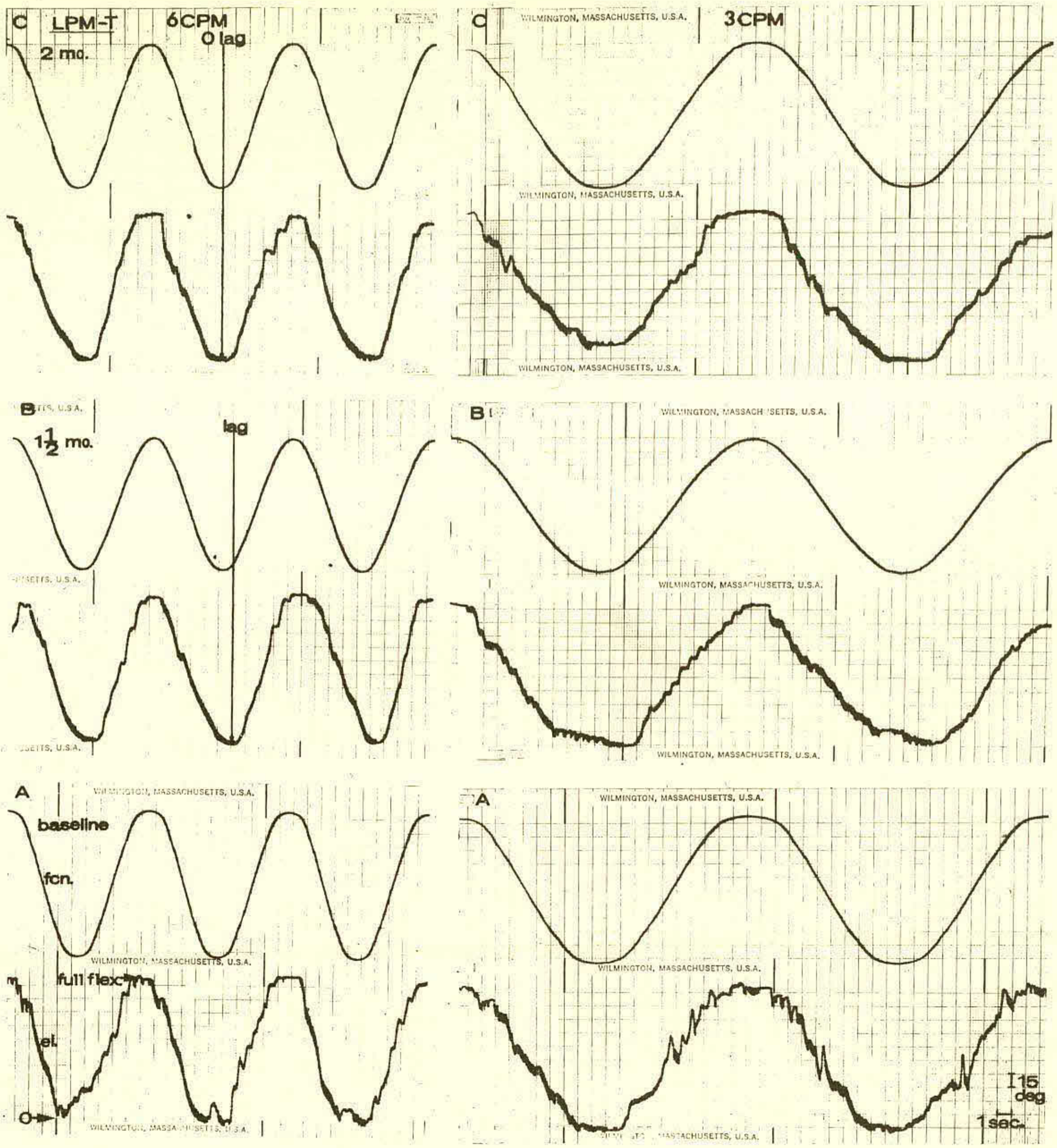


FIGURE 4. Comparison of performance on tracking version of limb position monitor at beginning, middle, and end of treatment.

of treatment. While the judgment as to what constitutes a "good quality" movement is admittedly subjective, the following criteria were used in making the evaluation: 1) overall shape (symmetry and amplitude) of the movement record; 2) presence or absence of extraneous reversals in the direction of movement; 3) presence or absence of tremor; and 4) presence or absence of overshoot. A movement was considered to be of good quality if its record was similar to the function generator output tracing in shape and amplitude, showed no extraneous reversals in direction of movement, was free of tremor (or showed reduction in tremor compared to the child's baseline records) and was free of overshoot (particularly in the case of the square-wave mode).

Figure 3 illustrates the progress made by a child using the original LPM device. In A, there are large scale oscillations (amplitude 50 degrees, frequency 2 per second) while the child moves his arm from one position to another, with an overshoot before settling down to the new position. However, the child can hold his arm reasonably steady at each new position. Note that as treatment continues within a single session, the tremor is reduced and stability is improved. In B, the child holds fixed positions well, with a shorter period of higher frequency oscillation during the transition from one position to another, but attempts at repetitive movement between 70 and 110 degrees of elbow flexion reveal irregularity. Also, tremor increases during these movements. After five additional treatments, in C, the tremor is almost gone, holding is good, and repetitive movements have a uniform, smoother profile. The child can make fine movements (10 degrees at a time) without tremor or overshoot. At the end of two month's treatment time, in D, the child has sufficient control to make either large- or small-scale movements with a minimum of tremor, can hold any position assumed, and can make repetitive movements of almost identical time course and amplitude. This child is ready to progress to use of the LPM-T device.

Figure 4 presents a comparison of stripchart records made using the LPM-T device at three different times during treatment of another child. Improvement is demonstrated in terms of better "shape" of individual movements as treatment progresses, with elimination of extraneous reversals in the direction of movement and reduction or elimination of tremor. Note that it is the faster movements (6 per minute), which are the best initially and show the greatest improvement, which is contrary to the common assumption that rapid movements are the most difficult for the athetoid cerebral palsied child to execute. (The fact that slow movements are the most difficult to control is consistent with the theory that the control problem is due to de-sensitized muscle stretch receptors; since the receptors are normally "tuned" to sense rapid movement rather than slow movement, when they are de-sensitized slow movements will be even farther out of their operating range.)

For all of the children using the LPM-T, the percentage of "good quality" movements was higher for 6 per minute tracking than for 3 per minute tracking, and it was higher for square wave movements at 6 per minute than for sinusoidal movements at 6 per minute. In the case of four of the children, the percentage of "good" movements was higher during the third period of treatment than in the first. The smallest gain was from 10 to 12%, and the largest from 4 to 33%. Obviously, there is still a lot of room for improvement with additional training.

3) Multiple Axis Limb and Head Position Monitor (MAD)

This device is comprised of five independent position indicators moving along parallel vertical tracks, each driven by its own servomotor. Each servomotor receives its input from a potentiometric movement transducer, and is provided with a potentiometric offset control in addition. The purpose is to provide the child with visual indicators of the positions of five body parts simultaneously, and thus to teach him to achieve and hold coordinated positions of particular functional significance. The primary goal is to teach the child the correct relative positions of the head, elbow, wrist, and fingers for self-feeding (the majority of CP children cannot feed themselves independently because of their poor coordination, and require the help of an aide or other personnel during feeding periods; this increases the cost of providing for them as well as hampering their self-reliance). In using the device, the offsets on all five channels will be adjusted so that when the child succeeds in aligning all of the indicators at the midpoints of the vertical tracks he will have positioned his head and arm correctly to enable self-feeding. Alignment of all indicators at the midpoints of their tracks is the end-point for all tasks; the offset settings will determine the postures thus achieved.

The device has been tested using the helmet from the Head Control System and the elbow and wrist sensors from the Limb Position Monitor to provide inputs; however, there is no provision as yet for recording from the device. In the final version, an additional transducer for finger position will be provided and there will be two timers, one to measure total treatment time and the other to measure the time during which all of the indicators have been held at the midpoints of the five tracks simultaneously. In this way, the percentage of the total treatment time during which a particular position was held can be determined. And, the coincidence detector for timing the holding of positions will also be used to run the movie projector, via the projector controller built for the Head Control System, as a reward for holding the correct position.

None of the children could be expected to be able to control all five channels simultaneously when they begin training with the Multiple Axis Device. They will practice with each indicator moving individually (that is, controlling one joint or one plane of movement of the head at a time) and then begin adding and combining channels as their skill improves. Once they have been able to reach and hold the correct position for feeding for at least 10 seconds, we will further reinforce their efforts by placing a spoon in their hand with a viscous food (such as honey) while they practice holding the position and eating. Simply being able to eat by themselves in this way should be the most powerful reinforcer conceivable for these children; this would represent a major step towards socialization for them.

Some preliminary studies which bear on the use of the MAD were done with the Head Control System and Limb Position Monitor in combination. Only one child had progressed enough to try this; the records in Figure 5 show the preliminary results. The task was more difficult than the one which would be presented by the MAD, since using the HCS and LPM together means the child must simultaneously use auditory cues for head position and visual cues for limb position (in fact, he probably time shares - monitoring head position for a short time, then switching over to limb position, then back to the head, etc.). The probability of performance deterioration due to information overload is great here, until independent control of each body part becomes automatic. This child has not reached that stage, but his efforts do result in short intervals of simultaneous stability of head and limb. At other times, head movement induces arm movement and an accidental arm movement leads to a subsequent head movement. Obviously, a child must learn to move body parts independently, to coordinate movements of various parts simultaneously, and to move various parts in particular sequences. We hope to accomplish this in training with the MAD, where the child can use visual cues for monitoring all body parts of concern and the indicators are close enough to each other to permit rapid scanning across them and thus better time-sharing during the monitoring process.

4) Visual Voice Monitor (VVM)

This device consists of a Tektronix D15 Storage Oscilloscope, two microphones, and a specially built plug-in unit which provides amplification and filtering for the microphone inputs. The purpose of the device is to provide a "voiceprint" or visible record of the sounds the child emits, and to use this visible record both to motivate the child to perform and provide him with feedback about the results of his attempts at correct sound production. Depending on the task chosen, the child can use one microphone to practice pre-speech activities, or the

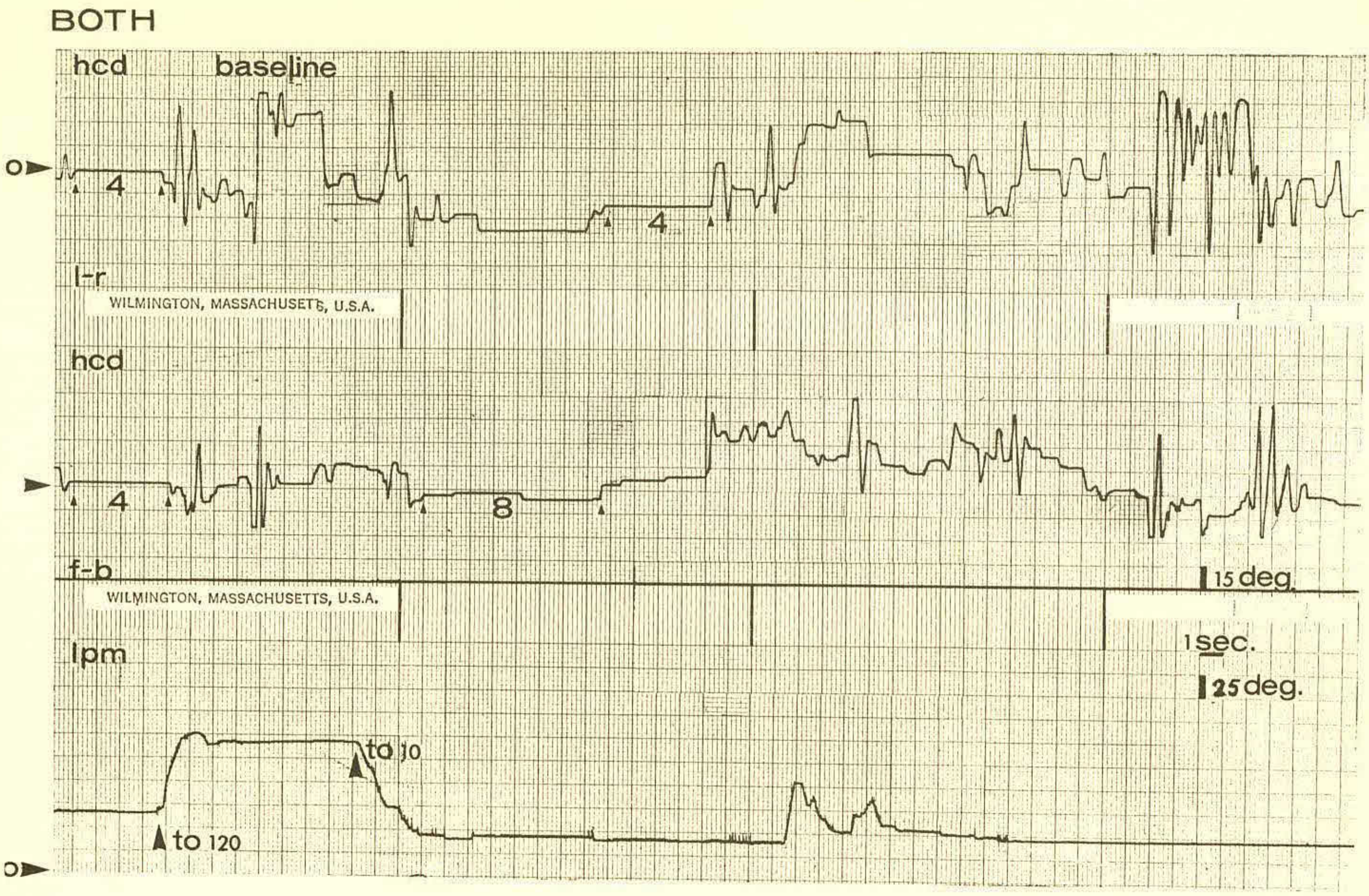


FIGURE 5. Stripchart recordings for head control system and limb position monitor used simultaneously (approximation of multiple axis limb and head position monitor).

therapist can speak into one microphone to set up a model for the child to match as the child speaks into the other. Pre-speech activities during the past year included working to sustain vowel sounds, working for breath control (increasing the number of sounds emitted per single expiration), and working for intensity modulation. Actual speech training activities included working for correct pronunciation of vowels and for pronunciation of beginning and final consonants. The filter settings and sweep speed were varied to suit the task at hand. The beginning of the sweep is triggered by hand or by the voice, depending on whether consonants or vowels are to be pronounced. Vowels are used for voice triggering since they yield repeating patterns (consisting of a fundamental frequency with harmonics superimposed) throughout the trace, while consonants are higher frequency transients which must be "caught" in mid-trace rather than used to start the sweep. A remote "erase" button is provided for the child or therapist to erase the traces after they have been studied and/or compared.

Four cerebral palsied children were involved in this portion of the program. Two of them worked to sustain vowel pronunciation (both being rather shallow breathers and initially lacking sufficient inspiratory capacity to sustain sounds on expiration). In both cases, the use of the VVM immediately prolonged sound production, adding several seconds to the duration obtained using a stopwatch for timing, as is customary. One of these children also worked for pronunciation of the final consonant "d" at the end of short words, and as soon as the VVM was introduced, the rate of pronunciation of this sound went from 0 to 100%. The "d" sound was not well blended with the rest of the word, however, and the therapist is working on this problem at present. Another child worked for intensity modulation. Initially his voice was completely "flat", but with the use of the device he is able to vary loudness considerably. The therapist is currently teaching him proper speech inflection patterns so that he can apply his new found ability to modulate his voice. The fourth child had presented such a behavioral problem that the therapist had almost given up hope for establishing rapport. He was fascinated with the VVM, however, and began to look forward to coming to speech therapy after it was introduced. The teacher established rapport with him through the intermediary of the device, and at this time, therapy is proceeding at an excellent pace - with the child operating the device almost exclusively! He has mastered all the switching and dial-setting operations necessary for "self-teaching", and the therapist simply establishes goals for him and helps him interpret the visual feedback while he practices.

5) Automated Visual Discrimination Apparatus (AVD)

This device is used for visual perceptual training, and is programmable to administer visual discrimination tasks ranging from simple (e.g., are two stimuli same or different) to complex (e.g., find two parts out of a choice of four possibilities which together make up a whole figure matching the test stimulus). To date, it has been used largely for a match-to-sample task with four choices being available to the child. That is, for each discrimination task, he looks at a test stimulus projected on a fixed panel at the far left, then scans across four movable panels to find the single figure which exactly matches the test item with respect to both form and orientation. He indicates his choice by pressing directly on the panel where he sees the figure, thus activating microswitches which position-code his response, and if he is correct, he is paid off with an M&M candy before the next slide in the sequence is shown. If he is wrong, there is no reward but the projector does advance to the next slide. Thus, the system operates on a no-correction basis.

The system consists of a commercially available (Lehigh Valley) modified Kodak Carrousel Slide Projector (the modification being a light-sensitive "reader" which determines the correct answer for each slide, which is coded on the slide itself), a projection box which contains solid-state circuitry for comparing the coded answers with the child's responses and has a front panel with five opaque sub-sections onto which the stimuli are projected, and an M&M candy dispenser (also Lehigh Valley).

The slide format for problem presentation is used to provide maximum flexibility of including and sequencing items according to the needs of individual students. The discrimination material on the slides has been prepared by a commercial artist and follows a sequence of increasing complexity dictated by the findings of Hubel and Wiesel concerning the neurophysiological mechanisms of visual perception (3). The child begins by matching straight line segments of various lengths and orientations, then moves on to matching angles of various sizes and orientations, then deals with curved lines, then with geometric figures composed of straight and curved lines, then with single capital letters, then with lower case letters, then with groups of letters, then with words, etc.

The projection box is equipped with two counters, one for total responses and the other for correct responses, to record the child's performance. In final form, the system will be interfaced with a teletype which will print out an item-by-item analysis of each test run (correct or error for each response, and time to respond) as well as the total correct.

To date, the system has been used with a group of 37 learning disability (dyslexic) youngsters and with 51 primarily physically handicapped children. Comparison of the performance of these two groups shows that many of the physically handicapped children have visual perceptual problems comparable to those of the dyslexic group. That is, while one's attention is first drawn to the obvious motor problems of the physically handicapped group, use of the visual discrimination system also calls attention to their having more subtle perceptual problems. Thus, these children are multiply handicapped, having sensory-perceptual problems as well as problems with posture and movement due to brain damage, and their educational program should meet both their needs in both these realms.

Use of the system also provides a method of determining the true intellectual potential of children who do not communicate well due to speech problems or who have limited motor response repertories. Several children whose appearance makes them "look" retarded got perfect scores on their visual perceptual test runs; this level of visual perceptual ability is indicative of higher intelligence than had heretofore been established with standardized IQ tests. When one considers the output limitations placed on these children by poor speech and lack of motor control, this is not too surprising. Since the AVD was designed so as to require only minimum coordination and effort for a child to "tell" his teacher what and how he sees, its application may lead to better intelligence assessments than have been possible with verbal and/or pencil-and-paper tests used to date.

The test scores obtained using the AVD have been correlated with the ages of the children, a process wich may provide another way to diagnose mental retardation. For most children, performance is related to age; the older the child, the fewer the discrimination errors they make. However, there is a group who are old enough to do better than they manage to score, even with repeated attempts. Their insufficiently developed visual perceptual skills may be indicative of general retardation.

The children enjoy using the AVD and are genuinely motivated to do well when using it. After all, this is a situation where giving correct answers really does "pay off" and you can get the answers right even if you cannot hold a pencil.

6) Gait Training Apparatus (GTA)

This device is used to teach the child to contact the floor with both his heel and his toe as he takes steps, to walk in correct gait patterns whether unassisted or using aids such as crutches, and to keep his toes pointed straight ahead rather than pointed in or out while he walks. Many CP children walk on their toes and toe in or out due to muscle imbalance, and this device is intended to inform them of their errors so that they can correct them.

The GTA presently consists of a battery-powered multiple tuned oscillator unit, pressure-sensitive switches, and headphones for both student and therapist. A switch is placed beneath the heel and toe of each shoe, and also at the pressure point of any assistive apparatus (crutch, cane, walker, etc.) used by the child in ambulation. There are three switches for each side of the body; closure of each switch results in production of one of the tones in a major triad. Activity of one side of the body generates tones an octave higher than the other. Thus, in using the device, the child "hears himself walk", and he is instructed to generate particular tonal sequences which correspond to desired gait patterns. The therapist monitors the child using the headphones (and visually as well) while he practices the desired patterns. The heel-strike, toe-strike sequence for each foot is of particular concern. The child can practice this walking in place as well as when he is traversing the room. He can also improve his dynamic balance, one of the requisites for independent walking, by practicing shifting his weight from one foot to the other, shifting weight from heel to toe, etc., while standing in place.

To date, the GTA has been used with only two children, and there is no provision for recording their progress other than filming. Both children were able to make heel contact with the floor upon trying the device for the first time, whereas without it they were "toe-walkers". (Supposedly toe-walking is due to involuntarily generated extensor hypertonicity, but clearly the child can voluntarily override plantar flexion at the ankle resulting from extensor hypertonicity if he is given a concrete goal to accomplish - namely, producing the tone corresponding to heel contact - and if he can directly monitor the results of his efforts.)

In the final version, a foot alignment detection feature will be provided based upon null-point seeking within an external electrical field, using auditory feedback, and telemetry will be used to mark the pattern of foot switch closures on a strip-chart recorder for quantitative studies.

7) Anti-Drooling Device (ADD)

Most CP children allow their mouths to remain open constantly, and thus drool rather than swallowing their saliva. This contributes to their "dull" facial expression and makes them unpleasant to work with in close-contact situations. The ADD was developed to teach the CP child to keep his lips closed, with the anticipation that the saliva will then be swallowed either voluntarily or due to the operation of the swallowing reflex which is triggered by the accumulation of fluid in the mouth.

The device consists of a timer unit, a pressure-sensitive switch, and a reinforcement delivery mechanism which can be used either with a coin dispenser or the projector controller originally developed for the head control system. In the initial stage of training, the child is rewarded by receipt of a token each time he keeps his lips pressed together for a time exceeding the duration set on the timer unit. A counter records each token delivery, and daily records of token receipts provide the basis for measuring learning. As the child's ability to control his lips increases, the duration for which he must maintain switch closure will be progressively increased to 99 seconds. When the child can keep his lips closed for this duration, the reward will be changed to watching movies. In this mode, the timer setting is reduced to a value which corresponds to the maximum amount of time during which he can allow his lips to relax and the switch to open before the movie will stop. Here, the number of interruptions during movie watching will be counted as an index of progress.

To date, the device has been used with only one child; he had progressed from a 5 second to a 10 second best duration lip closure within one week of training with the device. (At this stage we are uncertain whether an aspirator will have to be used along with this device as saliva accumulates in the closed mouth, or whether voluntary and/or reflex swallowing will occur as anticipated. If it does not, then aspiration of saliva will be used as a form of mild negative reinforcement if the child does not swallow within a selected, variable period of time after fluid buildup begins.)

8) Communications Device (COMM)

The latest addition to our program will be a communications system, still in development. This system will incorporate three major components: 1) an alphanumeric character display matrix, 2) a Telray alphanumeric character display CRO, and 3) a teletype machine. The system is intended for use by children lacking intelligible speech, and also will allow more accurate assessment of the intellectual capacity of the speech impaired CP child by circumventing his defective verbal output channel (obviously this bypass will only be temporary, since speech training must be provided and printed communication must not be allowed to substitute totally for verbal communication unless it is the last resort).

In operation of COMM, the characters in the display matrix will be illuminated in a sequential scan; the child will operate a switch when a character he wishes to select is lit, and that character will be displayed on the CRO and typed on the teletype simultaneously. Each child will be interfaced to the machine by switches appropriate for his motor capacity. In addition to choosing the characters, the child will also be able to control the scanning rate. And, after a message is spelled out he will be able to edit the CRO display so that a final error-free hard copy can be printed out from the edited teletype tape.

For maximum speed, the scan will be down the first column and then across the row which the child selects with his first switch closure. After each character selection, the scan will return to the starting point and begin over. For flexibility, the position and sequence of characters in the display will be variable, with changes accomplished through insertions of differently programmed ROM's into the circuitry. As a first step, we have separated the vowels and consonants, the former being presented first, and we have sequenced the characters according to their frequency of occurrence in words typically used by this particular population of individuals in the school setting. Numbers and math operators are presented last, so the children can use the lower portion of the display for doing math problems.

For preliminary training while the COMM is under construction, we have tried a few children with a device, marketed by Fairchild, which corresponds roughly to the display matrix portion of the system. They select letters, words, and phrases using a hand switch which steps a light one unit at a time in either vertical or horizontal direction across the columns and rows. There is no temporary storage or hard copy; the therapist records the message one letter or word at a time. However, the application of this device is limited to those children with sufficient manual dexterity to operate the hand switch, and that excludes a great many children. We intend to provide alternate switching manipulanda so that this commercially available device can be used in training for use of the COMM system or independently of it, since our one prototype device cannot possibly serve the needs of all the speech impaired children in the program. They will have to take turns in "saying their piece"; and, we are anticipating some "mind-blowing" statements from children who have never been able to communicate their thoughts before.

An incident along this line which illustrates the hidden intelligence of these children occurred when our program was demonstrated for the cerebral palsy unit of a state residential care facility. While one of the youngsters was trying the head control system for the first time and, with it, doing quite well at head stabilization, a staff member asked about the cost of the device. I stated the multi-hundred dollar figure and was about to argue the case for use of such technology in education despite the high initial cost. Before I could do so, another child who was a friend of the fellow using the head control device began speaking. His articulation was so poor that I could not understand him. The message interpreted by one of the therapists familiar with his speech problems was our best endorsement to date. He said, "It's worth it!"

References

1. Asanuma, H., "Cerebral Cortical Control of Movement," 1973, The Physiologist, Vol. 15, #3: 143-166.

2. Harris, F. A., Spelman, F. A., and Hymer, J. W., "Therapy for Cerebral Palsy Employing Artificial Sense Organs for Alternatives to Priopriceptive Feedback," 1972, Proceedings of the 1972 Carnahan Conference on Electronic Prosthetics, 1-4, Office of Research and Engineering Services, University of Kentucky.

3. Hubel, D., and Wiesel, T., "Receptive Fields of Single Neurons in the Cat's Striate Cortex," 1959, J. Physiol., 148:574-791.

TYPEWRITER FOR TEACHING SEVERELY HANDICAPPED CHILDREN

by

P.J.R. Harding*, Y.J. Kingma+, A.J. Kozak*
K.A. Stromsmoe+ and J.F. Vaneldik+

* K and H Engineering Ltd
Edmonton, Alberta,
Canada.

+ Elect. Eng. Department,
University of Alberta,
Edmonton, Alberta, Canada.

Summary. A compact, portable Electronic Typewriter Controller has been developed to be used in conjunction with a standard IBM Model 72 electric typewriter. The completed system allows a severely handicapped child to type using only the simple actions that he is capable of performing. Three modes of operation permit flexibility in the design of the patient interface unit so as to best utilize individual abilities.

Introduction

The inability of many severely handicapped children to communicate effectively with their teachers makes learning very difficult. Children suffering from a spastic or paralyzed condition often find speech impossible, so a system which allows such a child to type can provide the necessary feedback between the student and the teacher.

Several typewriter systems have been developed for these handicapped people such as POSSUM(1) and VOTEM (2). At the University of Alberta, two working models were developed and used by the Glenrose Provincial Hospital (3,4). Their successful operation led to a request for a third typewriter. Rather than duplicate the previous circuitry it was decided to take advantage of the most up-to-date electronic products available and develop an entirely new design.

The physical operation of the Electronic Typewriter Controller is basically the same as other previous systems. A simple operation such as moving an arm; kicking a foot, actuates a switch. The first part of the switching sequence is translated into a position on the lightboard. The next movement causes the desired symbol to be typed.

The layout of the display lightboard corresponds identically to a typewriter keyboard. This results in a reduction in possible typing speed, when compared to an optimum layout (4,5). However the advantages of the system are a considerable reduction in cost and size, and the easy implementation of another control mode, that is joy-stick manipulation.

System Description

The overall system consists of a standard IBM Model 72 electric typewriter and an Electronic Typewriter Controller. The Controller is made up of three sub-assemblies; a lightboard monitoring unit; a control and actuator unit; and an interface unit. A block diagram and a photograph of the completed assembly are shown in Fig. 1 and Fig. 2 respectively.

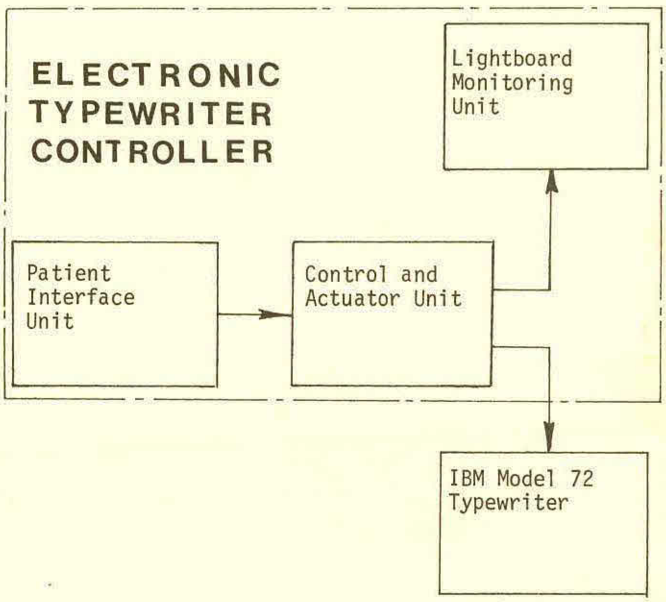


Figure 1 Block Diagram of Electronic Typewriter Controller

IBM Model 72 Typewriter

Initially the possibility of using an electronically controlled typewriter such as a Teletype was considered. The Controller would have generated the required electronic signals. Long delivery times and high initial cost ruled out this approach. A standard electric typewriter, on the other hand requires that the typewriter keys be actuated by a mechanical movement. Commercially available solenoids can be modified to perform this task.

The IBM Model 72 typewriter was chosen because its feature of interchangeable type balls gave it added flexibility. Using solenoids to activate the keys, no modifications were necessary to this machine. Thus by simply detaching the Controller, the typewriter can be used in its original configuration.

Lightboard Monitoring Unit

Below each symbol, a light emitting diode (LED) is located on the monitoring unit shown in Fig. 3,

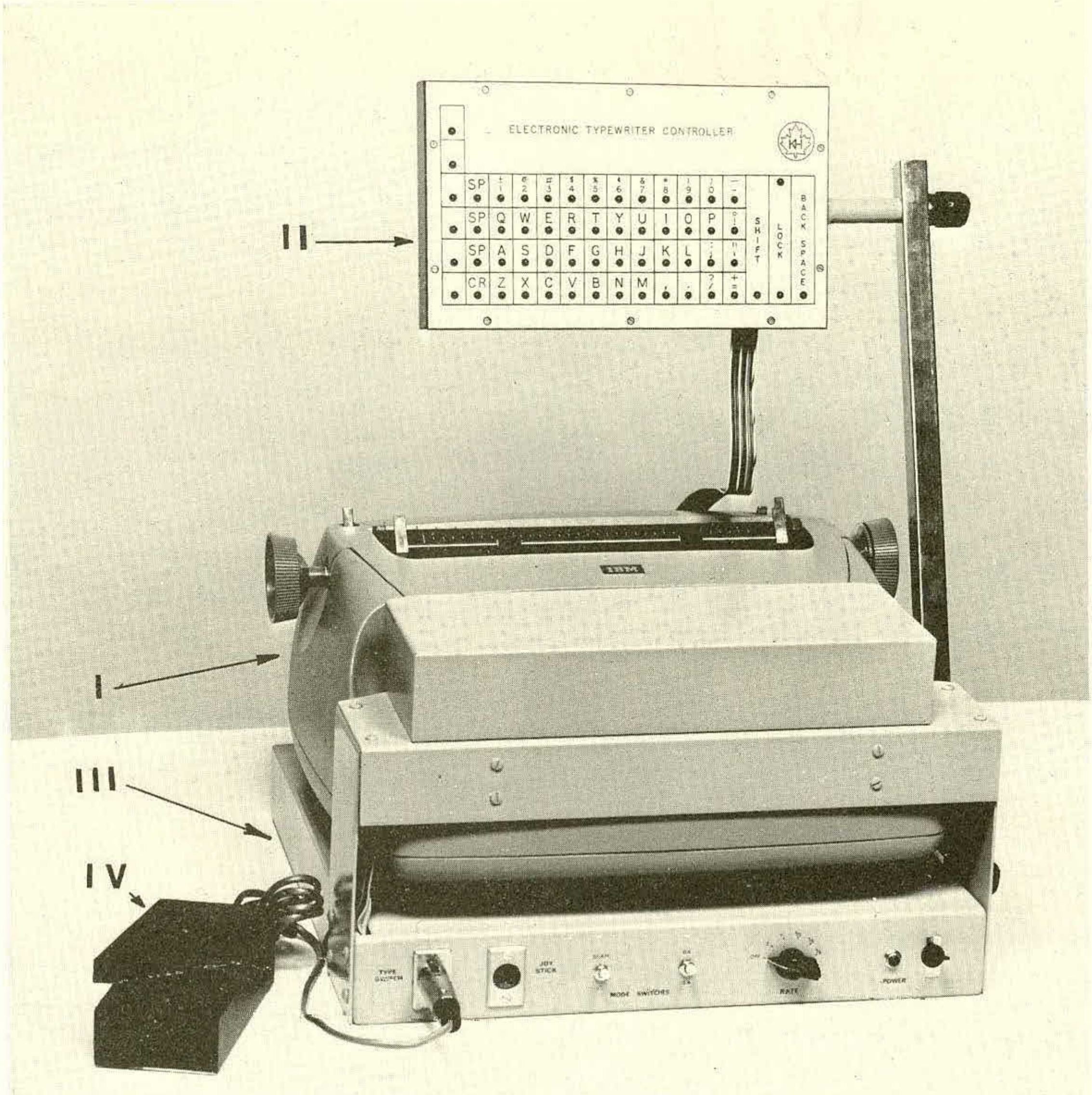


Figure 2 Electronic Typewriter
Controller (I) IBM Model 72 electric typewriter,
(II) Lightboard Monitoring Unit
(III) Control and Actuator Unit (IV) Interface Unit

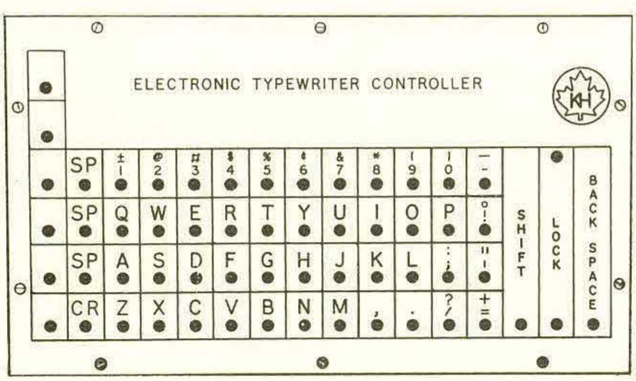


Figure 3 Lightboard Monitoring Unit

which indicates both the status of the control and the symbol ready to be typed. There are fifty-eight lights, for in addition to the forty-four typewriter symbols, the display board has positions to initiate space, shift, lock, carriage return and back space. Even if the system were to be used for a different language, only the display board panel would need to be changed to correspond to that country's basic typewriter keyboard.

The display unit is situated above the electric typewriter and the handicapped person can observe both the typed letters and the lightboard simultaneously. The position of the monitoring unit can be easily adjusted to individual preferences. When a LED is on below a specific symbol, the patient is able to type that character by initiating the proper switching action.

Control and Actuator Unit

The control and actuator unit takes the electronic signals from the patient controlled interface unit and converts them into the desired typing commands. Two power supplies, the logic circuitry and the solenoid drivers for the unit are located in a chassis below the typewriter. The forty-nine solenoids are situated in a box above the typewriter keyboard. These two parts, which make up the control and actuator unit, are indicated in the photograph of Fig. 2.

Located on the front of the lower chassis of the unit are the ON-OFF switch, two mode control switches, a five position clock speed control switch, and connector plugs for the interface unit. The two power supplies, the logic control circuitry and the solenoid drivers each make up a printed circuit board which is inside the chassis. The fuses and power supply cord are at the back of this assembly. Two cables also lead from this part of the control unit; one to the lightboard to show the status of the control unit and the second to the solenoid box.

Interface Unit

A great deal of time and effort must go into the evaluation of the individual capabilities of each handicapped typewriter recipient to assure that the best switching method willbe used to activate the interface unit. The initial interface system was relatively easy to construct as it was designed for a patient able to operate a switch with a downward motion of one arm.

Operating Modes

The controller has been designed to operate with either a single acting (SA) or a double acting (DA) switch and for either SCAN or JOY-STICK operation. In the SCAN mode, the initial starting position of a sequence is the upper left hand corner of the lightboard. This is referred to as the "home" position. Upon closing a switch, the LED directly below the "home" position lights up, (called the "start" position). With a speed determined by the clock, the control moves in sequence through the first vertical column on the left hand side of the display board. When the operator initiates a second closing of the switch, the control starts moving sequentially through the positions in the desired horizontal row. The third switching action causes the control to stop at a specific position and a fourth closing of the switch (within three clock periods) results in a type pulse being sent to the indicated solenoid. After the type pulse, the control automatically returns to "home" position.

With the exception of the "lock" position, only one LED can be turned on at a time. The LED's in the "start" position and the other four positions in the first column are not associated with any typewriter functions, they act as a buffer to allow sufficient patient reaction time between selection of a row and selection of a character in the adjacent column. The "lock" position shifts the typewriter into upper case, and the typewriter will remain in the locked position until a "shift" pulse is initiated. The upper LED in the "lock" location of the lightboard remains on while the typewriter is in the upper case position.

The SCAN mode of this machine requires more patient initiated switching actions than some others, for example the TUFTS Iteractive Communication (6). The advantage of always starting from a specific initial point ("home") is that the same time sequence is required to type each given letter.

In the DA mode each change in a switch's position has the same effect as the momentary closing of a switch in the SA mode. This type of operation is desirable for patients operating a pressure switch with mouth controlled "suck and blow" action.

If the operator does not initiate a switching action while the control is scanning the first column, it will continue to cycle through this column's lower four positions, omitting only "home" and "start". Similarly, if the control is scanning a horizontal row, it will continue to cycle in that row until a switching signal is received. In this way, if the typist's reaction is slow initially, he can simply wait for the row or column to recycle without starting the sequence over again.

In the JOY-STICK mode, the control position is moved either right-left, or up-down. A separate switch is used to initiate a type pulse. The "joy-stick" may be either a conventional aircraft type joy-stick for an individual with relatively fine control or it could be a quad of four switches for a person with less control. Neither the "home" or "start" positions are used in the JOY-STICK mode.

The clock can be set to give timing pulses with a period of 2, 1, 1/2, 1/4 or 1/8 second. In the SCAN mode, only the longer periods are feasible, whereas in the JOY-STICK mode the faster times are suitable.

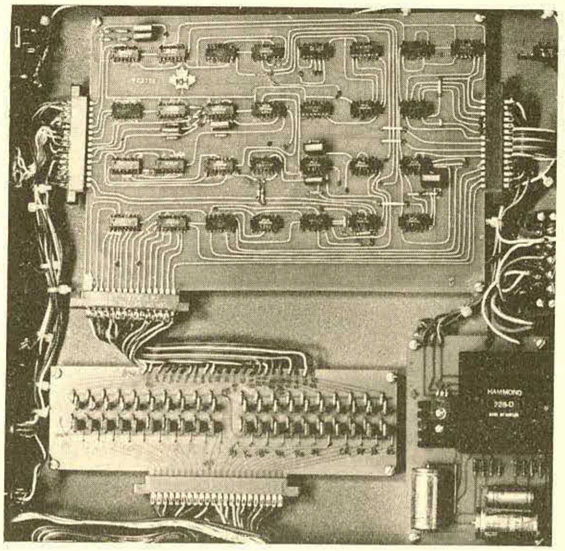


Figure 4 Power Supplies, control logic and solenoid drivers of control and actuator unit

Construction

In addition to the use of LED's in the display panel, CMOS logic is used throughout. The two power sources are a 28 volt unregulated supply to drive the solenoids and an 8 volt supply for the logic circuitry and LEDs. All electronic components are located on removable printed circuit board cards (Fig. 4).

With the lightboard and the typewriter keyboard having the same physical layout, an x-y grid control can be used to actuate both a LED and the corresponding solenoid. This feature greatly reduces the number of solenoid drivers required and simplifies the interconnection wiring. The net result is a compact unit that is readily portable.

Conclusion

The present typewriter has been designed to minimize total costs and reduce physical size. The simple structure and the use of solid state components results in a highly reliable system. The flexibility offered by the various control modes permits a typewriter interface unit to be designed for almost all handicapped people.

References

(1) Maling, R.G. and Clarkson, D.C., "Electronic controls for the telraplegic (POSSUM)" Paraplegia, 1, 3, 1963 pg. 162-174.

(2) Newell, A.F. and Nabaui, C.D., "VOTEM: The Voice Operated Typewriter Employing Morse Code", Journal of Scientific Instruments, (Journal of Physics E.) Ser. 2, Vol. 2, 1969.

(3) Kildaw, R, A Multi-terminal Interface Allowing Typewriter Operation by Paraplegics (MINTOP), E.E. Dept., University of Alberta, M.Sc. Thesis, 1968.

(4) Englehart, T.W., A Computerized Typing System for the Handicapped, E.E. Dept. University of Alberta, M.Sc. Thesis, 1971.

(5) Kohn, D. Codebreakers Macmillan Co, London, England, 1967. pp. 100-101.

(6) Foulds, R.A., "The Tufts Interactive Communicator", Proceedings 1972 Carnahan Conference on Electronic Prosthetics, Dec. 1972, pp. 16-24.

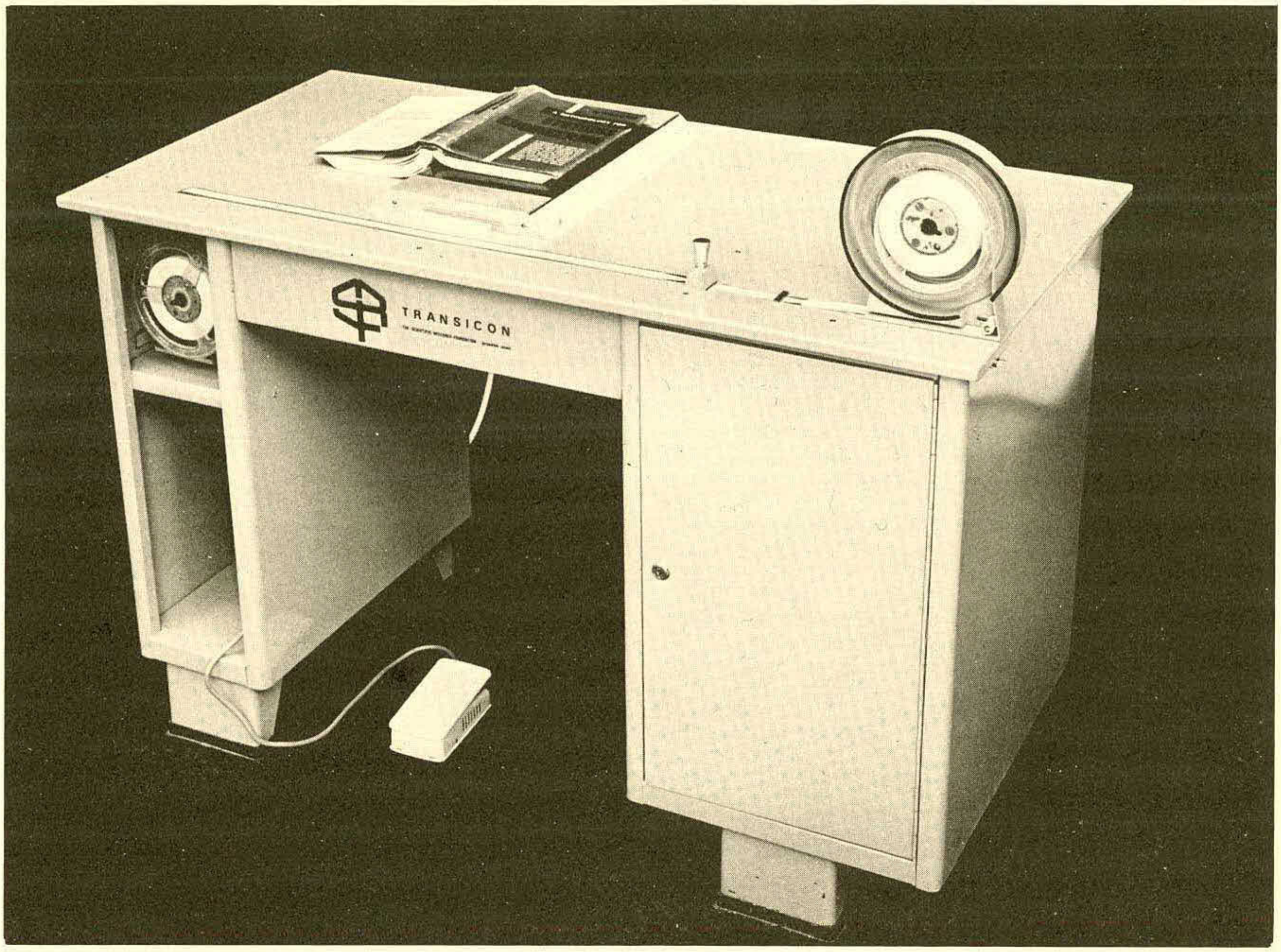


Figure 1. The TRANSICON.

out the first line; locking on to it and tracking it even if the book is somewhat tilted. Depressing the second button enables the automatic magnification to work, while preventing the head from advancing to the next line. When this appears correctly on the Braille tape, one of the LEAD buttons is depressed, locking out the automatic magnification circuit and allowing the head to advance down the book, one line at a time. The Braille tape is stored as it is embossed in a loop of about 3-4 lines and the reader advances it at a rate convenient for him by means of a knee-control.

When the head reaches the bottom of the page, the LEAD button is automatically released and the head returns to the top of the page. The reader must turn to the next page manually.

3. Principle of Operation

The reading head projects an image of the line of print on a

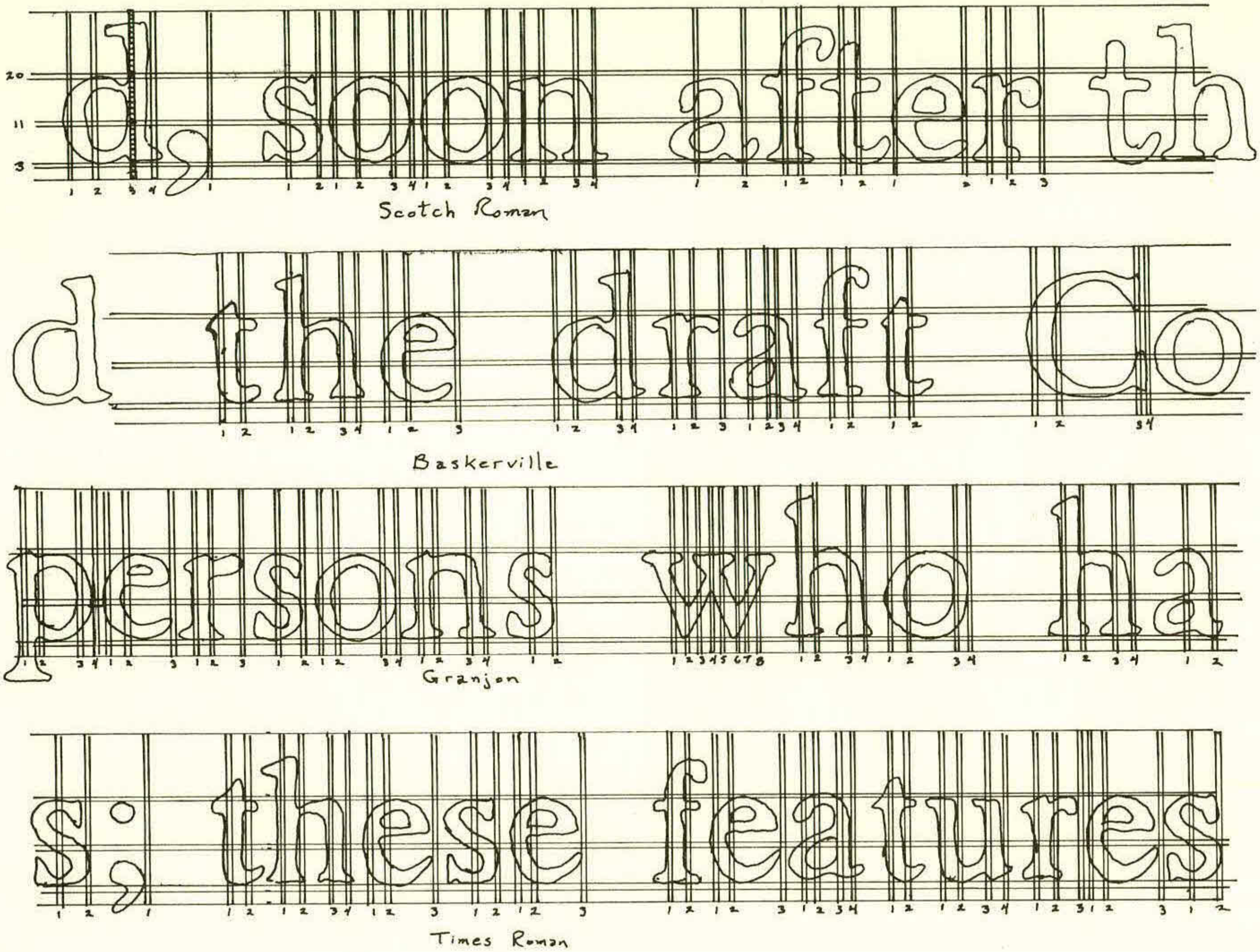


Figure 2. Typical printed lines as they appear on the detector.

64 element detector which is vertical to the line of print and sweeps along horizontally, as shown in Fig. 2. The 64 detectors are combined into 32 detector pairs such that if either is black the combination indicates black. The primary reason for this is that many features of interest, such as the bar on the letter e or the neck on the letter m may be extremely thin and require "thickening" for positive identification. In addition, while it is generally <u>not</u> necessary to detect a white space between adjacent characters, the accuracy of reading is improved if this can be recognized, especially in cases of multiple joining, that is, where several characters would otherwise appear to be joined together. By using detectors only 50 microns (2 mils) wide, the joining of extremely close characters is reduced.

Two detectors, identified as nos. 3 and 20, are designated to skim the bottoms and tops of the lower case letters. This is accomplished by smoothing the difference in their outputs, and using this difference to move the carriage up or down. Remembering that each of these consists of an adjacent pair connected so that either black indicates black, it will be seen that in principle, this difference control is capable of keeping the head aligned to the print within an accuracy of 1/2 detector; e.g. 3% of the height of the lower-case letters.

This operation is predicated on the size of the characters being exactly as shown. In practice, this is achieved by using the <u>sum</u> of detectors 3 and 20 to increase or decrease the magnification until this condition is very closely approximated.

It is apparent that the above alignment criterion is sensitive to the appearance of ascenders or descenders in the line of print. However, the vast majority of ascenders and descenders cross vertically through the zone traversed by detectors 3 and 20, so that when these detectors are black, so are their lower and upper neighbors, nos. 2 and 21. When this occurs, we weaken the contribution of the alignment detectors to one-half. This value was selected as adequate to achieve good alignment, while still allowing the head to align quickly in the case of gross misalignment. The dynamic properties of this alignment will be discussed below in section 5.2.

Since we are required to read books and other documents having varying widths, it is impractical to set right and left margins as in the case in other readers. Instead, we determine the line ends by deciding that the detector has scanned an all-white region for a predetermined time and using this information to reverse the direction of the carriage. If no line is present, the head will reach the edge of the page, where it is reversed by either the limit switches or by going off the page. In the latter case, the detector appears to see an all-black region; thereby reversing the carriage.

The scanner thus sees a row of vertically-normalized characters passing by. For multi-font operation, these must now be horizontally-normalized (See Fig. 3). The difference between the characters of many fonts is often the ratio of the horizontal to the vertical size, e.g., an m, or an n may vary by as much as 40% from one font to another, thereby making 'template' matching techniques difficult. In the TRANSICON horizontal normalization is achieved as follows: First, one detector, no. 11, is designated as the SYNCH detector. The output of the detector is delayed and whenever there is a change in sign all the detectors in the array are sampled.

An additional sample is taken when all detectors first see white.

The outputs of the 32 detectors are then combined in certain OR combinations, so that if any detector is black, the combination is black. There are 12 such combinations in the TRANSICON and since a maximum of four samples are used per character, there are a total of 48 positions stored in the shift register memories.

This synchronous sampling technique results in the loss of valuable information which may exist between samples. For example, the break in the serifs of the letter h may be extremely useful to separate it from b. This type of information is retained by using latches which can detect whether certain features were ever present (e.g. if the character was an ascender or a descender), or if certain features were ever present between two specified samples as in the h above. This technique has been extended to detect features which may occur even before the first sample; e.g., if no lower serif has appeared before the first sample, the character is a t and not r. In general, this technique is applied to a number of detectors together, a particularly striking example of which is the detection of the bar on the letter e. In this case, we look at all of the detectors between nos. 9 and 15 simultaneously. Only if all of them are ever white, is it a c. This never occurs in an e.

Finally, we combine the information from shift registers, latches and the number of the sample to decide which character is being scanned. Detection is always on the last sample from the SYNCH detector, or, in a very few cases, on the all white sample. The reason for this is that, due to the high incidence of joining, especially in poorly printed books, it is necessary to set all of the shift registers and latches either at the start of a character (first black), or when the SYNCH detector goes black.

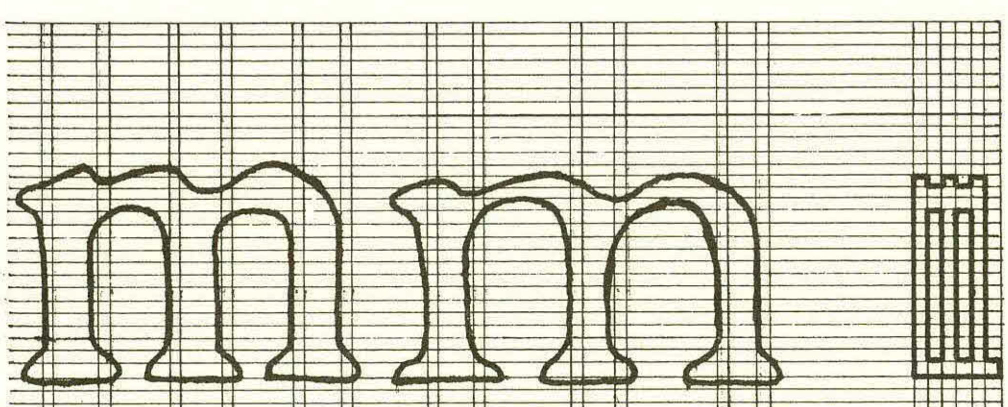


Figure 3. Horizontal normalization.

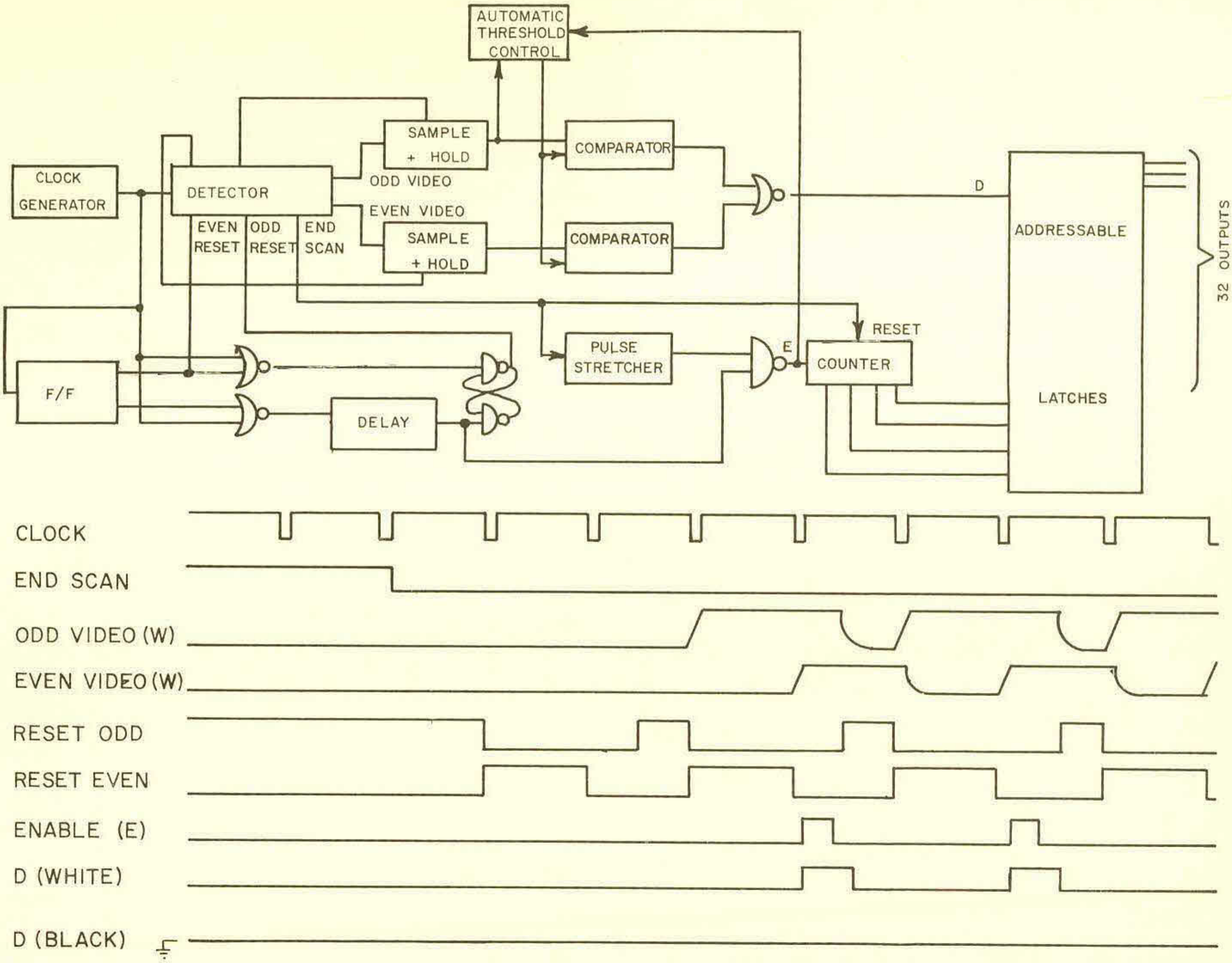


Figure 4. Detector electronics (schematic).

The characters are converted into the six-bit Braille code and stored in a one-bit-wide buffer memory before being used to emboss the tape. This is because certain characters may occur faster than the Brailler can work. The Braille-embossed tape is stored in an intermediate loop at the machine reading speed; from where it is removed across the reading area at a speed determined by the reader. If the reader falls behind by 3-4 lines, the machine will stop and wait for him to catch up.

4. Detector Electronics.

Fig. 4 shows the block diagram of the detector electronics. The detector is a self-scanning linear array of 64 silicon photodiodes, operating at a clock rate of 68 khz. The output consists of two video trains; one for the odd numbered detectors and one for the

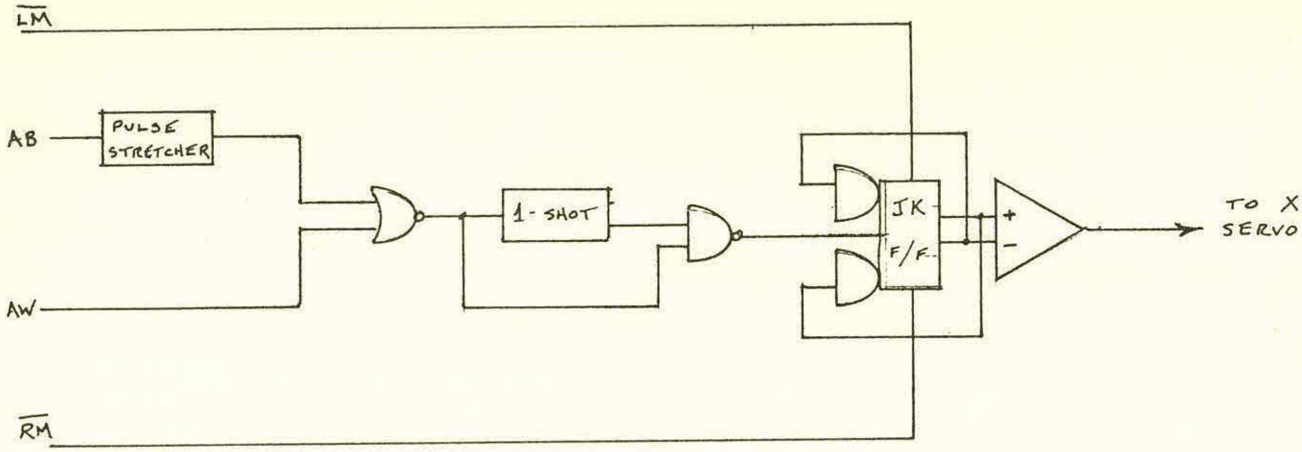


Figure 5. X controller.

even ones. These pulses are used to charge integrating capacitors which are reset after the information is utilized. Since we wish to combine the information from adjacent pairs in an OR gate, the voltage on the odd video capacitor is retained until the even video capacitor is charged. Each of these voltages are compared to the threshhold voltage and the outputs of the two comparators are then logically OR-red. The threshold voltage is derived by averaging the positive and negative peaks from one of the video lines. These indicate the whiteness of the paper and the blackness of the ink respectively. Thus, the threshold voltage is automatically set to 50% of the way from the white to the black, regardless of the actual values.

5. Controller

5.1 X Controller (Right-Left)

Fig. 5 shows the block diagram of the X controller. The control information is generated when the detectors are all white or all black.

In general, the head scans in either direction at the same speed until it reaches one of the mechanical limits. These are reed switches operated by a pair of magnets located on the carriage. If the carriage reaches one of these, the flip-flop F-1 is set to the opposite direction. If, however, the all-white appears for a time corresponding to approximately three characters, the direction of the carriage is reversed. If no line is present, the carriage will continue to the edge of the page where the all-black will reverse the direction. The machine begins to read automatically at the start of a line and at the end jumps down to the following one. In the case of a short line, as is often the case at the end of a paragraph, the machine must scan the new line for a minimum of five seconds before going into read. This prevents misalignment errors, particularly if there is some error in setting the LEAD.

5.2 Y Controller (UP DOWN)

See Fig. 6. There are several sources of information for the Y motion of the controller; the line following, the LEAD, the Rapid Advance and the BT command, which returns the carriage to the top of the page. These will be discussed in detail below.

First of all, as mentioned previously, the line following action is achieved by using the average difference between two detectors which skim the bottoms and tops of the lower case letters. The outputs of these detectors are weakened by their lower and upper neighbors respectively to reduce the effects of ascenders and descenders. Since this error is integrated before being applied to the Y axis follow-up servo, the resulting action would be quite

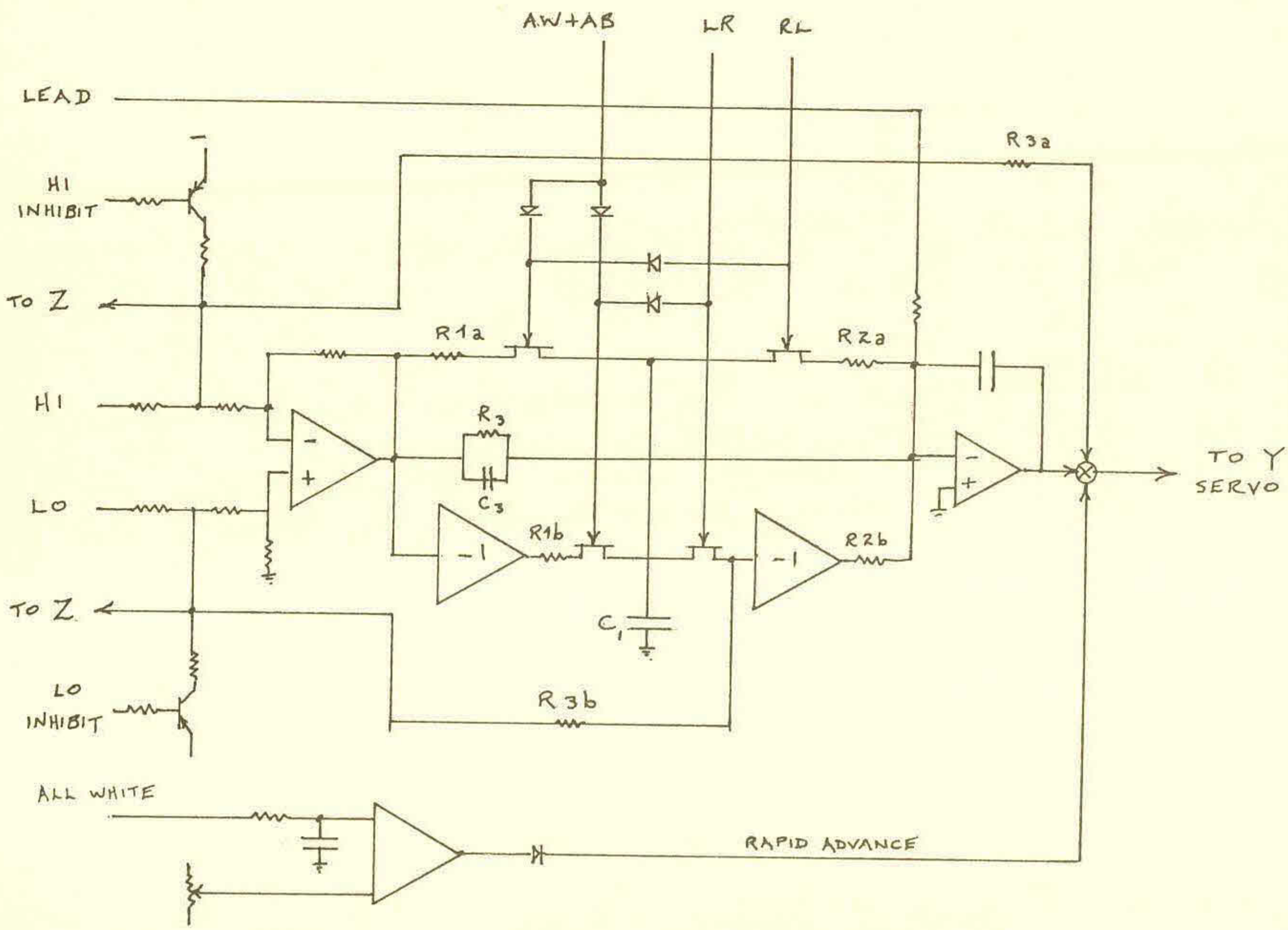
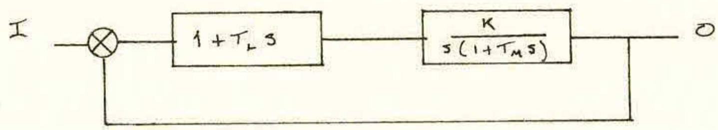


Figure 6. Y controller.

Follow-Up Servo:



$$\frac{O}{I} = \frac{1 + T_L s}{1 + (T_L + \frac{1}{K})s + \frac{T_M}{K} s^2}$$

Closing the loop with:

$$\frac{G}{s(1 + T_A s)},$$

The root-locus becomes:

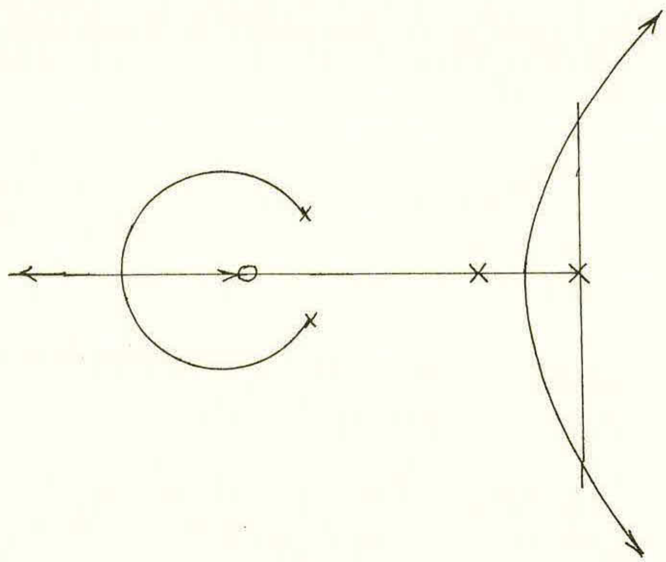


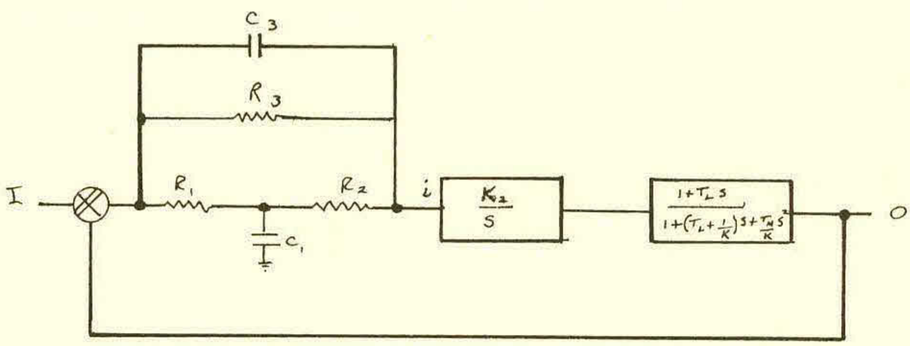
Figure 6A. Uncompensated line-following.

unstable, as shown in the root-locus plot in Fig. 6A. In this diagram, the follow up servo is indicated as having a tandem compensation network, which leads to the closed-loop transfer function shown.

The addition of resistor R 3 and capacitor C 3 adds the two central zeroes to the root locus plot as shown in Fig. 6 B. Note the inherently high stability even at high gain. The addition of R 3 alone is adequate for stability, but leads to poor transient response.

Tracking is improved by allowing the voltage on the averaging capacitor to float during intervals in which no information is present; e.g. all-white or all-black. In addition, special care is taken to enable the lines to be tracked even in the presence of tilt. This is accomplished by switching the sign of the error signal both before and after the capacitor according to the direction of travel of the carriage. Thus, the voltage on the capacitor is proportional to the tilt of the line and does not change in sign when the carriage reverses.

The LEAD command is applied to the carriage each time that a line has been read, provided one of the LEAD switches is depressed. The present model provides for four preselected values of LEAD; 11,



$$G = K_2\frac{\left\{R_1+R_2+R_3+[(R_1+R_2)R_3C_3]s+R_2R_3C_1C_3s^2\right\}(1+T_L s)}{R_3(R_1+R_2+R_2C_1s)\left[1+(T_L+\frac{1}{K})s+\frac{T_M}{K}s^2\right]s}$$

The root-locus becomes :

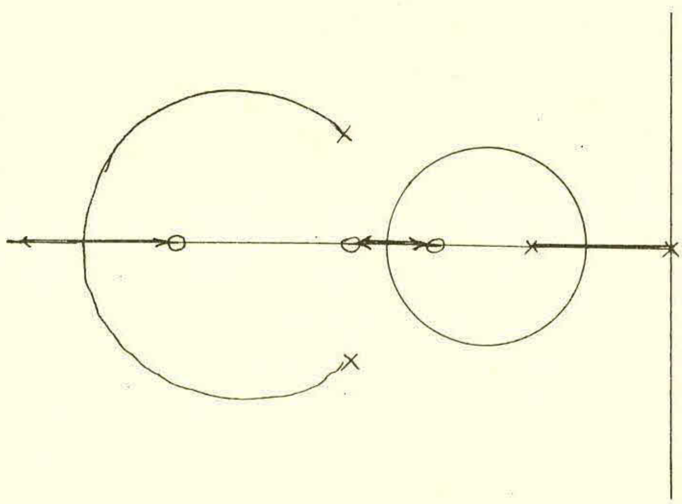


Figure 6B. Compensated line-following.

12, 13 and 14 points. If the wrong switch is depressed, the carriage will correct the error on the flyback, provided that the line is sufficiently long. In the case of a short line an interlock is provided to prevent the machine from going into read until it has scanned the short line several times.

If no print is encountered for one full scan, a signal is generated which causes the carriage to advance rapidly until a line is encountered. This allows the machine to proceed through blank areas relatively quickly.

Finally, at the bottom of the page, a mechanical means if provided to release the LEAD button; the head automatically returning to the top of the page. The reader must, of course, turn the page manually.

5.3 Z Controller (Magnification).

The sum of the voltages from the control detectors 3 and 20 is a measure of the relative size of the characters imaged on the detector. In this case, the averaging time constant is somewhat longer, and the weakening effect of detector 21 is reduced to 75%. This is to prevent control-reversal when an extremely large change in magnification is called for. The magnification is optimized for 9 and 11 point type, but retains adequate focus over the range from 8 to 12 point. When one of these limits is reached, the magnification circuit is interlocked to prevent damage to the unit.

6. Synchronizer

The circuit shown in Fig. 7 introduces a delay in the output of detector 11 and generates a strobe pulse each time there is a change in sign. The delay circuit also separates strobe pulses which would otherwise occur too fast. Thus, a white gap in the middle of a v which may be shorter than the delay time will generate one sample after one delay interval and a second one delay interval after the end of the gap. A short black region on the other hand (one that is shorter than the delay) will generate one sample at the end of the black region and a second one delay interval later. These waveforms are also shown in Fig. 7. Since the appearance of the all white is used to generate an additional strobe, and since this has no delay (in order to reduce joining effects), there could be some uncertainty in the strobing at the end of a letter such as an o. We prevent this by including the delayed D 11 in the all white circuit rather than D 11 itself.

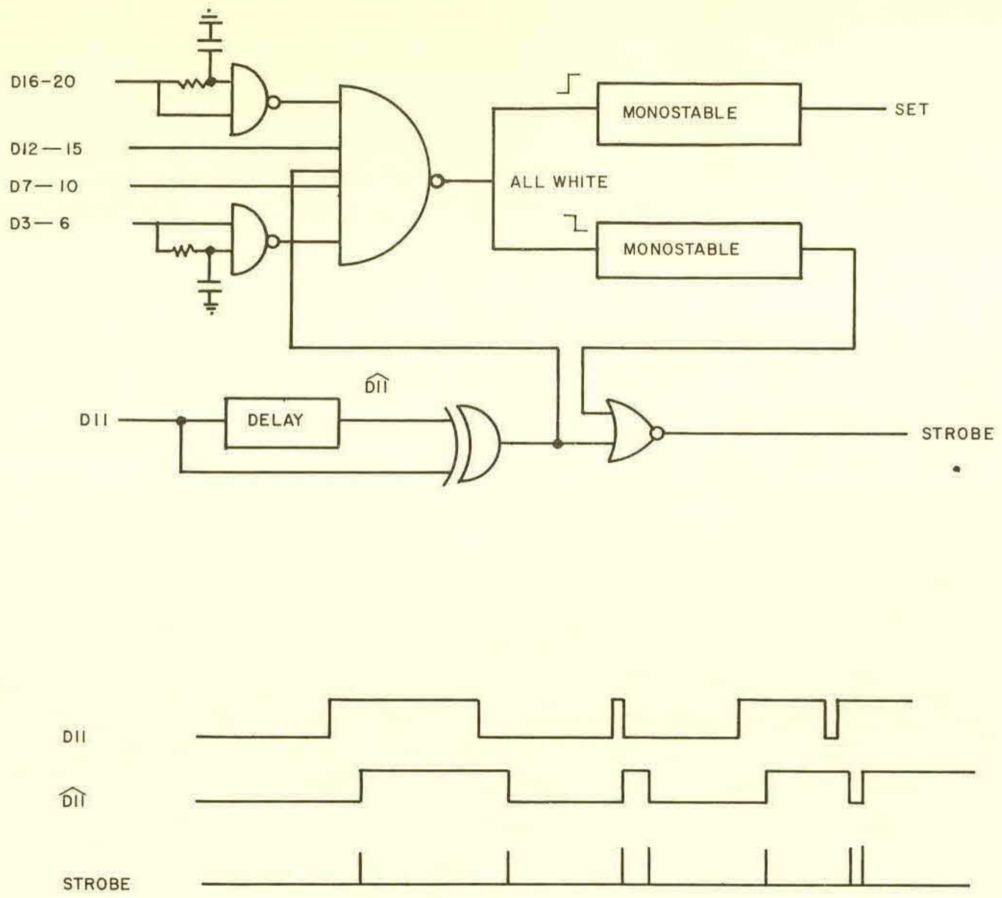


Figure 7. Synchronizer.

Often two adjacent characters, while not actually touching, will not be separated by a vertical line of white. This is especially true of an f, which often overlaps the following letter. Since, however, the all white circuit includes only detectors from just below the baseline to just above the body line (lower case letters), this problem is avoided. However, overlapping of lower case letters is dealt with in the TRANSICON with an additional means. Each of the detectors in the all-white lying near the baseline or body line is examined in an OR gate to determine whether it is now white, or was recently white. If either is true, its part of the all white is activated, thus resulting in a kind of 'snake circuit' for the detection of the all white.

7. STOP READ (see Fig. 8)

Before proceeding to the description of the logic, I should like to mention briefly the operation of the stop read circuit. This prevents recognition of an additional character once one character has been

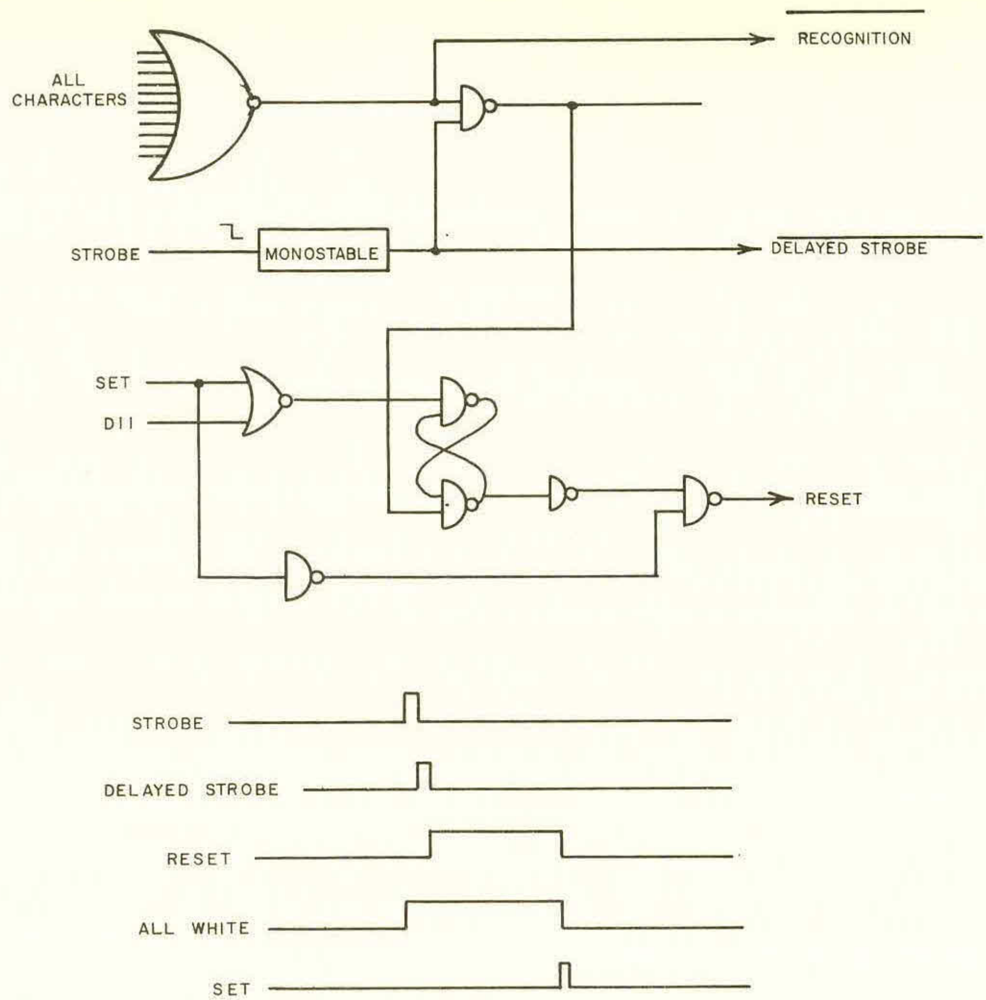


Figure 8. Stop read.

recognized. At the start of the strobe pulse, the shift registers enter the information present on the detector combinations. Following this, a delayed strobe pulse is generated which operates the counter and enables the logic to recognize a character. If a character is recognized on this strobe, a Recognition signal is generated which, at the end of the delayed strobe, operates the RESET latch. This latch is returned to the SET condition (the RESET is lifted) by one of two means:

One is the first black command (end of the all white), which is required if the following character is a., e.g. the synch detector never sees it and the only information is generated on the all white strobe.

The second is when D 11 is black, which occurs in all the rest of the characters. This condition permits recognition of a new character even in the case of joining. If, however, no character was recognized at all, the RESET is supplied by the first black. Finally, if a condition should occur whereby the reset command occurs while D 11 is black, the memories will nevertheless be reset correctly.

8. The Latches

We mentioned these briefly in the general description. There are two types of latches; general and count-connected. These in turn are black-sensing and white sensing; thus, a black-sensing latch is useful to decide whether a character is an ascender or descender, while a white sensing latch is useful to detect the gap in the letter i, independent of its actual height.

Table 1 lists all the latches and gives an example of the use of each.

9. The Shift Registers

The twelve four-bit shift registers sample data from one or more detectors each. The detectors sampled by each are listed in Table 2 below together with the listing of the positions. The last information in carries the first digit 1.

TABLE 1 -Latches

Latch	Detectors	Criterion	Example
L 1	1	Ever Black	Descenders
L 2	3 - 10	Ever White	t (against f)
L 3	5 + 6	Either always white	?
L 4	8 - 10	Ever Black	.
L 5	16 - 20	Ever White	against f
L 6	19 - 22	One always white	i (against l)
L 7	21	Ever Black	t
L 8	23 - 29	Ever Black	Ascenders
L 9	9 - 15	Ever White	c (against e)
L 10	3 - 10	Ever White on P 2	h (against b)
L 11	8 - 10	Ever Black on P 2	E (against l)
L 12	22	Ever Black on P 1	against Y
L 13	23 - 29	Ever White on P 2	U

10. Logic

Information from the shift registered, latches and counter are combined to generate 75 characters according to the logic shown in Table 2 below.

TABLE 2

Shift Register	Detectors	Positions			
1	1	101	201	301	401
2	3 - 4	102	202	302	402
3	5 - 6	103	203	303	403
4	7	107	207	307	407
5	9-10	105	205	305	405
6	12 - 13	106	206	306	406
7	13	107	207	307	407
8	15	108	208	308	408
9	16 - 17	109	209	309	409
10	18	110	210	310	410
11	19 - 20	111	211	311	411
12	23 - 29	112	212	312	412

SEARCH AND RETRIEVAL TECHNIQUES FOR USE WITH TAPE RECORDINGS

by

Emerson Foulke, Ph.D.
Perceptual Alternatives Laboratory
University of Louisville
Louisville, Ky. 40208

The development of speech recording technology has resulted in a significant increase in reading effectiveness for blind persons. However, no adequate way of enhancing the recorded format with the index information needed for efficient search and retrieval has yet been developed. Two related approaches to the solution of this problem are under investigation in the Perceptual Alternatives Laboratory. In one approach, low frequency tones are mixed with speech signal during recording. When a tape so recorded is reproduced in the fast-forward mode of operation, these tones are heard as high-pitched "beeps" that are clearly audible and easily differentiated from the background of chatter resulting from the reproduction of the recorded speech signal at the fast-forward tape speed. In the other approach, speech is used to present index information. Spoken index information is recorded at a speed that approximates tape speed during fast-forward operation. The index tape thus prepared is reproduced at the normal playing speed and its signal, now inaudibly low, is mixed with the signal on a tape containing the full text, and this combined signal is copied on the final master from which the duplicates used by listeners are generated. When the listener's tape is played at the fast-forward tape speed, the spoken index information is reproduced at a speed that is close enough to the speed during recording for it to be intelligible. As in the case of the tones in the index code, it is displayed against a background of chatter.

A Comparative Evaluation of Simple Index Codes for the Identification of Locations in Tape Recorded Texts

Background

A serious problem experienced by those who read by listening to tape recorded texts is the difficulty of finding specific locations, such as chapter headings, paragraph headings, or the locations at which new pages begin in the inkprint analogues of such texts. Because these locations cannot be found easily, the retrieval of desired information is an inefficient process, and as a consequence, tape recorded texts are less effective than their inkprint analogs.

The retrieval problem may be reduced by recording, at desired locations in the recorded text, and on the same track with the recorded text, tones so low in frequency that they are nearly inaudible. Because they are nearly inaudible, when the tape is played back at the speed used during recording, their presence does not interfere with the perception of the speech recorded on that track. However, when the tape is played back in the fast-forward mode on a tape recorder that has been modified so that the tape remains in contact with the playback head during fast-forward operation, the tones are increased in frequency by an amount proportional to the increase in tape speed, and they are heard as clearly audible "beeps," displayed against the background of high-pitched chatter that results from reproducing the recorded text at the fast-forward tape speed.

This system of indexing is currently employed by Recording for the Blind in its preparation of recorded textbooks. A rudimentary code with two characters is in use. One character, a single "beep," signifies the beginning of a new page. The other character, which is composed of two "beeps" of the same type, signifies the beginning of a new chapter. An announcement at the beginning of each track indicates the pages covered on that track, and if a new chapter starts on the track, this fact is also announced. With this information, the reader can, by interpreting the code signals that are manifest during fast-forward operation, locate pages and chapters in which he is interested.

Rationale for the Composition of Index Codes

An index code with a larger number of characters should permit a more detailed search of a tape track, and the result should be more efficient retrieval. There are many ways in which additional characters for such a code may be created, but constraints that should be observed in their composition are suggested by consideration of three factors--the time consumed by the presentation of code characters, the latency of identifications of code characters, and the ease with which the code can be learned.

The time required for the reproduction of code characters should be kept at a minimum because, at the high tape speed at which they are reproduced, the precision with which desired locations are indicated declines rapidly with increased duration of code characters. In view of the small number of characters that would be required in an index code, the time consumed by the reproduction of a character can be kept within practical limits by restricting the number of elements that may be used in composing a character. The rule that will be followed in constructing the one or more codes to be tested is that a code character must contain either one or two elements. The two elements in a character may be the same, as exemplified by the two identical "beeps" used by Recording for the Blind to indicate chapter headings; they may be two different values in a single dimension, as exemplified by the dot and the dash in Morse code; or two elements may be produced by concommitant variation in two dimensions, for instance, a dot at one frequency and a dash at another; but a code character may not include more than two elements.

The number of identifications of different values in a stimulus dimension that can be made by a human observer is small. Furthermore, latency of identification increases as the values requiring identification are increased in number. In an index code of the sort under discussion, characters must be identified promptly, since the tape is running at a high speed during the time spent on identification, and the reader must be able to react quickly if he is to bring the tape to a stop in the close vicinity of the location marked by code characters.

An index code must be easily learned, since many of those who are candidates for its use will not receive formal training under the supervision of an instructor. If they learn the code at all, they will learn it on their own, under casually arranged conditions that may be inefficient for learning. Consequently, differences among elements must be highly discriminable, and the set of characters composed with these elements must be readily differentiated.

Tones are desirable as index code elements because their electronic generation is a simple matter, because they can be recorded satisfactorily, and because they can be varied in several dimensions in which human observers are known to demonstrate good discrimination. The one or more codes to be tested will include tones of two durations, tones of two pitches, and two portamento effects.

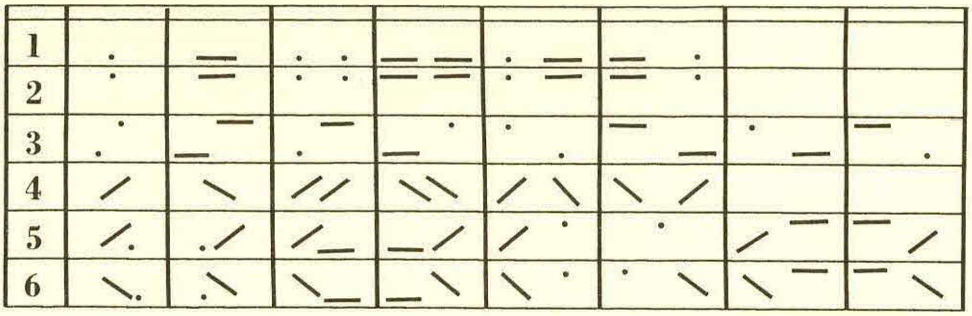
Two tones, differing in pitch by approximately one whole step in the musical scale, will be employed. It is virtually certain that any prospective user of an index code will have had an abundance of experience in detecting differences in pitch of this magnitude. The lower tone will be recorded at a frequency of 55 cycles per second, which is first octave A on the musical scale. The higher tone will be recorded at a frequency of approximately 61.735 cycles per second, which is first octave B on the musical scale. When reproduced at the fast-forward speed, the lower tone will have a frequency in the neighborhood of 880.00 cycles per second, or fifth octave A on the musical scale, and the higher tone will have a frequency in the neighborhood of 987.77 cycles per second, or fifth octave B on the musical scale.

Either of these tones may be presented at either of two durations. The shorter tone will be recorded for 480 milliseconds, and the longer tone will be recorded for 1,440 milliseconds. When these tones are reproduced at the fast-forward speed, the shorter tone will have a duration of approximately 30 milliseconds, the estimated value of a dot in Morse code sent at the rate of 30 words per minute, and the longer tone will have a duration of approximately 90 milliseconds, the estimated value of a dash in Morse code sent at the same rate.

One of the portamento effects will be a rising tone that increases in pitch by one whole step on the musical scale during its course. The other portamento effect will be a falling tone that decreases in pitch one whole step on the musical scale during its course. The two portamento effects will vary between the two values of pitch selected for use in the pitch dimension, and their duration will match the longer of the two durations at which tones of fixed frequency are presented.

TABLE 1

POSSIBLE COMBINATIONS OF DURATION, PITCH, AND PORTAMENTO



CODE:
- • = low pitched short tone
- — = low pitched long tone
- ˙ = high pitched short tone
- ¯ = high pitched long tone
- / = rising tone
- \ = falling tone

Table 1 shows the code characters that could be composed by exploiting every possible combination of duration, pitch, and portamento. This table contains 42 characters. Six of the characters are composed with only low pitched dots and dashes (Row 1), and six of the characters with only high pitched dots and dashes (Row 2). Since tape speed is not controlled by the capstan when a tape recorder is in its fast-forward mode of operation, tape speed gradually increases as tape is transferred from the supply reel to the takeup reel. Consequently, tones recorded at the same pitch will gradually rise in pitch if the tape on which they are recorded is reproduced in the fast-forward mode of operation. For this reason, a reader listening to code characters of the sort shown in the table will be unable to interpret code characters requiring absolute identification of pitch. He will be able to interpret characters requiring only relative identification; that is, he will be able to interpret characters in which both pitches are represented, because he can judge one pitch against the other. However, if the reader understands that when he hears a character involving only one pitch, it is to be identified as the lower of the two pitches, the six low-pitched characters can be retained, and only the six high-pitched characters need be sacrificed. Under this rule, referred to hereafter as the exclusion rule, the table contains 36 unambiguous characters.

A considerable amount of time and effort would be required to learn a code with 36 characters, and it does not seem likely that there would ever be a need for an index code with 36 characters. The task, therefore, is to select a set of characters that is large enough to make the discriminations required for indexing, but small enough to be learned easily. This selection should be guided by information concerning the ease with which characters of different types can be learned, and the speed and accuracy with which they can be identified after learning.

Table 2 presents nine codes that can be composed, using the dimensions of duration, pitch, and type of portamento effect, singly and in combination. The nine codes include three pairs of codes, the members of which are equivalent, and thus there are six different codes. The characters in parentheses are those disqualified by the exclusion rule. In the two- and three-dimensional codes, the characters surrounded by quotation marks are those in which not all of the permissible code dimensions are represented.

It is not necessary for all nine codes to be learned in order to obtain the desired information concerning the speed and accuracy of identification of code characters. There are five one-dimensional codes. All of the characters in one of these codes (1B) are disqualified by the exclusion rule, and this code can therefore be eliminated from further consideration. Two out of six characters are disqualified in two of the one-dimensional codes

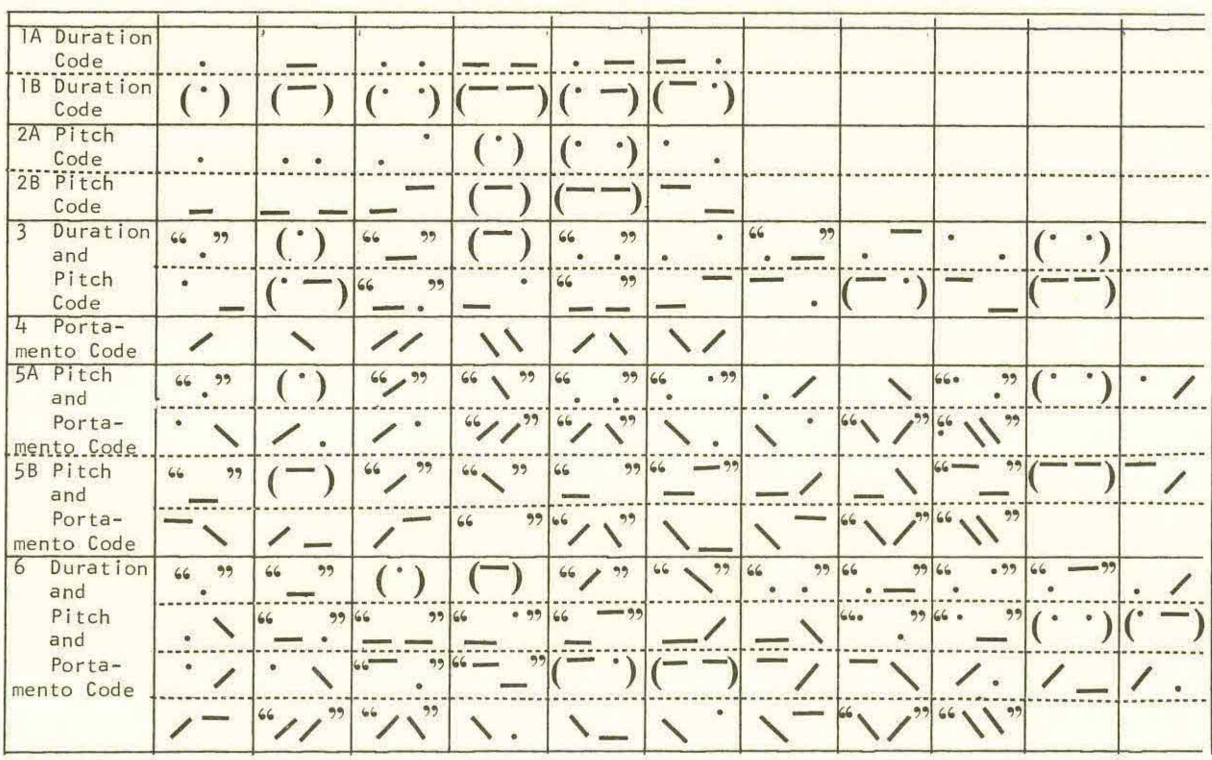
(2A and 2B) because of the exclusion rule, and since a code with only four characters would be inadequate for indexing, these codes may be dismissed from further consideration. All six of the characters in two of the one-dimensional (1A and 4) are permissible, and although it would be desirable to have more than six characters, an index code with only six characters would be useful. Therefore, these codes will be evaluated.

There are 20 characters in the code based on duration and pitch. However, six of these characters are disqualified by the exclusion rule. The 14 characters remaining should be ample for an index code, and therefore, this code will also be evaluated. However, in evaluating it, the six characters surrounded by quotation marks in Section 3 of Table 2 need not be considered. Since their identification requires the listener to decide only about duration, they are the same as the six characters in the duration code, already proposed for evaluation. The identification of the eight characters in Section 3 that are not enclosed by either parentheses or quotation marks requires the listener to decide about both pitch and duration, and these are the characters that must be evaluated in order to obtain the needed information about a code whose characters are defined by duration and pitch.

The two codes based on portamento and pitch are different only with respect to the elements that can be varied in pitch. In one case, those elements have the duration of dots in the duration code. In the other case, they have the duration of dashes in the duration code. The two codes should be approximately equal in terms of ease of learning. However, since less time is required to present the characters that include dots than to present the corresponding characters that include dashes, less time may be required to identify the characters that include dots, and since rapid identification is a requirement for precise determination of the locations marked by index code characters, the code with characters containing dots (5A) is the one chosen for evaluation. The two characters in this code enclosed by parentheses are disqualified by the exclusion rule. The eight characters enclosed by quotation marks involve only one of the two dimensions employed in this code. They also appear as

TABLE 2

INDEX CODES DEFINED BY DURATION, PITCH, AND PORTAMENTO



CODE: . = low pitched short tone; — = low pitched long tone; ˙ = high pitched short tone; ¯ = high pitched long tone; / = rising tone; \ = falling tone characters in the two one-dimensional codes proposed for evaluation, and need not be evaluated again. The identification of the 10 characters not enclosed by either parentheses or quotation marks requires the listener to decide about both pitch and type of portamento effect, and these are the characters that must therefore be evaluated.

The code based on duration, pitch, and portamento (Code 6) contains 42 characters, and includes all of the preceding codes as subsets. The six characters in this code distinguished by parentheses are disqualified by the exclusion rule. Either one or two of the three code dimensions are not represented in the 20 characters distinguished by quotation marks, and therefore, these characters appear in earlier codes already proposed for evaluation. In order to identify the remaining 16 characters, those not enclosed by either parentheses or quotation marks, the listener must decide about duration, pitch, and type of portamento effect, and these are therefore the characters that should be evaluated.

To summarize, five codes have been proposed for evaluation. Two of the codes contain six characters (Codes 1A and 4) and, in each case, all six characters need to be evaluated. One of the codes contains 14 characters (Code 3), of which eight characters need to be evaluated. One code contains 18 characters (Code 5A), of which 10 characters need to be evaluated. The final code contains 36 characters (Code 6), of which 16 characters need to be evaluated.

Method

Apparatus. A piece of apparatus has been constructed that generates all of the elements required for the composition of the code characters to be evaluated. This apparatus can be programmed by the setting of switches so that when a button is pressed, the elements required for a given character are generated in the proper sequence, and with the proper interval between elements. Controls on the apparatus permit the adjustment of the pitch and duration of tones, and of the inter-element interval. To prepare experimental materials, the output of this generator is recorded on tape. An oral reader can use the apparatus to add index information to the tape he is recording. In this case, the apparatus and his microphone are connected through a mixer to the input of his tape recorder. He must first program the apparatus for the code character that will be required next. When, in the course of reading, he arrives at the location that is to be indexed, he merely presses the button that initiates the generation of the programmed code character and continues reading.

The plan for evaluation. Five codes are to be taught to five comparable groups of college students, with 20 subjects per group. The learning tasks can be equated for difficulty by using all six characters in the two codes with only six characters, and by randomly selecting six characters for use from the supply of characters offered by each of the remaining three codes. The characters in each code will be given names, such as letters or numbers, and their identification will be learned by a paired-associates method in which, upon hearing a code character, the subject responds by pronouncing the name of that character. Following this guess, the experimenter pronounces the correct name of the character, thus affording the subject immediate knowledge of results. A trial consists of one random permutation of the six characters in a code, and trials are administered until a criterion of two successive errorless trials is attained. Each subject is then given 100% overlearning, so that all subjects will be asymptotic, or nearly so, with respect to speed and accuracy of identification. Finally, five trials are administered to each subject without knowledge of results, and responses are scored for accuracy and latency of identification. With this procedure, it will be possible to compare groups of subjects, and hence codes, in terms of the number of trials to criterion, the number of correct responses on the final five postlearning trials, and the median latencies of identification on the final five postlearning trials.

Using the information provided by these comparisons, one of the five codes will be selected for further investigation. To validate this selection, the code that has been chosen will be taught, in its entirety, to another group of 20 college students. As before, performance will be examined in terms of the number of trials to criterion, the number of correct identifications on postlearning trials, and the median latencies of identification on postlearning trials. If the outcome of this experiment does not clearly validate the choice that has been made, it may be desirable to construct confusion matrices and to evaluate each character in terms of stimulus ambiguity and response equivocation, in order to eliminate troublesome characters. If too many characters are lost by this analysis it may be necessary to subject the next-best code to the validation procedure.

When a code has been finally selected, it will be taught to a group of blind high school students. A tape recorded text will be indexed with this code, and the blind students who have learned it will be tested for accuracy of location, precision of location on first attempt, and the time required for location.

Voice Indexing of Recorded Texts

Just as index tones may be mixed with spoken texts and recorded on the same track, speech itself may be used to provide index information. Of course, if index information can be conveyed by recorded speech, more sophisticated indexing can be provided and the aural reader is not required to learn a special code. One application of such indexing is exemplified by the Talking Dictionary, the development of which has been undertaken by the Perceptual Alternatives Laboratory in cooperation with the American Printing House for the Blind and the Bureau of Services for the Blind, Kentucky Department of Rehabilitation Services.

The Talking Dictionary

The objective of this project is to record, on tape, the oral reading of a dictionary in a manner that will permit rapid search of the recorded format and efficient retrieval. The dictionary is presented on cassette. This cassette is reproduced on a slightly modified cassette player, but the modifications are simple and inexpensive, and properly modified cassette players are now available from the Library of Congress and the American Printing House for the Blind.

Method

One tape, referred to hereafter as the text tape, contains the full text of the dictionary. The other tape, referred to hereafter as the index tape, contains only the pronunciations of the words that are pronounced, spelled, and defined on the text tape. The index tape is recorded at 15 inches per second (ips), which is an approximation of the fast-forward tape speed of the modified cassette player on which the recorded dictionary is reproduced.

In order to prepare a dictionary cassette, the text tape is reproduced at the speed at which is was recorded, and the index tape is reproduced at 15/16 ips. The signals from these tapes are mixed and recorded on a final master tape that is suitable for use on a reel-to-reel duplicator.

To reproduce the dictionary cassette properly, two modifications of the conventional cassette recorder are required. Its motor speed must be continuously variable through a range that centers on a playback speed of 15/16 ips, and the playback head must not be retracted from the cassette when the machine is placed in the fast-forward mode of operation. This second modification permits the tape to be scanned by the playback head when the cassette player is running at the fast-forward tape speed. The cassette players now supplied to blind and physically handicapped readers by the Library of Congress, and the cassette recorders sold to blind and physically handicapped readers by the American Printing House for the Blind include the modifications required for the proper reproduction of the Talking Dictionary.

When a dictionary cassette is played on the modified cassette player at 15/16 ips, the signal containing the full text of the dictionary is reproduced properly. The signal containing only the pronunciations of words that are also spelled and defined on the full text tape is reproduced at a speed that is so much slower than the speed at which it was recorded that it is nearly inaudible. Only an occasional low-pitched rumble is heard and it is not seriously intrusive. When the cassette player is placed in the fast-forward mode of operation, the signal containing the pronunciations of words is intelligibly reproduced and displayed against the background of high-pitched chatter that results from reproducing the signal containing the full text at the fast-forward tape speed. Since at this speed the signal containing the full text is unintelligible, it adds only noise to the signal that is being processed by the listener, and it does not offer serious interference. As the listener scans at the fast-forward speed, tape speed gradually increases as tape is transferred from the supply reel to the takeup reel and the listener must make slight adjustments in motor speed in order to maintain the index signal at the proper pitch. Each index word is so located on the tape that its termination and the termination of the definition preceding the definition to which the index word refers occur at the same point on the tape.

Each cassette contains a braille label and a large print label indicating the first and last word recorded on each track. By referring to these labels, the listener selects the cassette and the track containing the word in which he is interested, and plays the cassette at the fast-forward speed until he hears the word for which he is searching. He then changes to the slow playback speed and listens to the spelling and definition of that word. Of course, the time required by a listener to consult this dictionary is considerably greater than the time required by the normal visual reader to consult a printed dictionary. However, if the alternative is a braille dictionary, the advantages of the Talking Dictionary are more apparent. The Talking Dictionary will require considerably less space and be significantly cheaper than the braille dictionary. Though a superior braille reader will consume less time in consulting a braille dictionary than in consulting the Talking Dictionary, the difference is not so marked as in the case of the visual reader, and it may disappear altogether in the case of a poor braille reader. Furthermore, most of the blind people who read by listening do not read braille at all, and for them the Talking Dictionary may provide the only alternative to dependence on the assistance of a visual reader in consulting a dictionary.

An automatic control devise has been constructed to operate the tape recorders on which the text and index tapes are reproduced. As the reader's voice is recorded on one track of a stereo tape, a tone from an audio frequency oscillator is recorded on the other track. After pronouncing each word on the index tape, and after completing each segment on the text tape, the reader presses a button which briefly interrupts the tone generated by the oscillator. When these tapes are reproduced, the recorded tones are sensed by circuitry which operates relays that start and stop the transports of the machines on which the tapes are being reproduced. When both tones are present or when neither tone is present, both transports run. If the tone on one tape is absent, the transport handling that tape is stopped and remains stopped until the termination of the tone on the other tape is sensed. Then the idle transport is started again, and the cycle is repeated. In order to make the termination of an index word coincident with the termination of the segment of text preceding the segment of text to which the index word refers, the tapes are played backwards.

<u>Other Applications</u>

Of course, if speech can be used to index locations in a dictionary, it can be used to index locations in other texts as well, and when the dictionary project has been concluded, additonal applications for this method will be investigated. For instance, it would be a simple matter to index a tape recorded text book with announcements that locate and identify new pages, new chapters, paragraph headings, and so forth.

In a project undertaken jointly by the Perceptual Alternatives Laboratory and Ada-Max Audio Visual, a design engineering firm in the audio-visual field, with the support of a grant from The Seeing Eye Foundation, a running abstract that is concurrent with the full text to which it refers, but only one-eighth as long, will be recorded on the index track for proper reproduction at the fast-forward tape speed. The aural reader will be able to scan the text recorded on this tape by listening to the abstract. When he encounters subject matter in which he is interested, he can change to the slow playback speed, and listen to the full text. As a ready example of this application, consider the manner in which a magazine such as <u>Newsweek</u> is read by most people. Rarely is it read word for word, cover to cover. Instead, the typical reader scans it rapidly, and slows down for careful reading of only those sections in which he is particularly interested. The provision of a concurrent abstract on the index track would permit the aural reader to read a recorded magazine in much the same manner.

THE STEREOTONER READING AID FOR THE BLIND, A PROGRESS REPORT

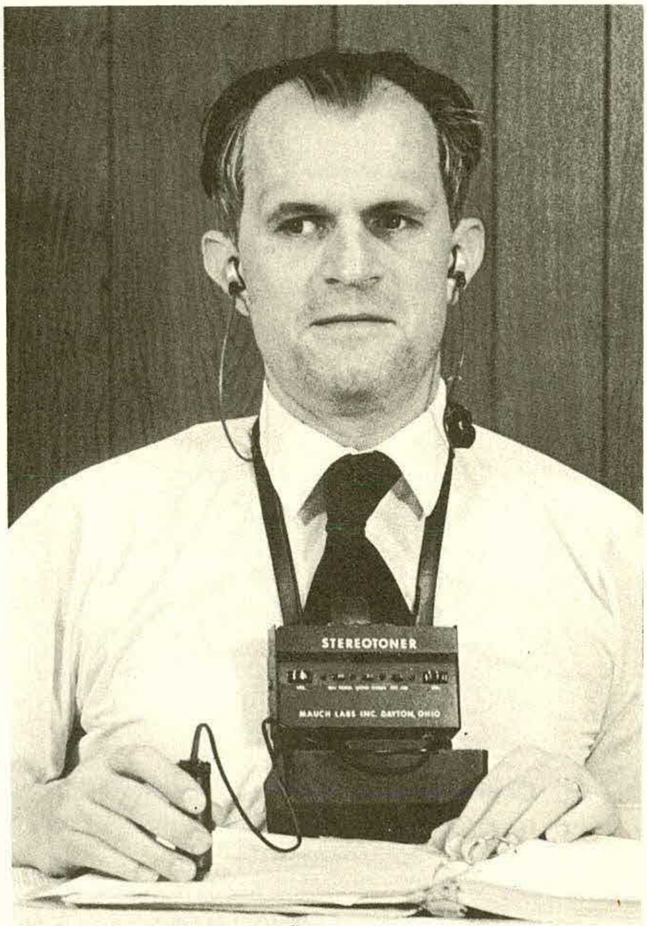
by

Glendon C. Smith
Mauch Laboratories, Inc.
3035 Dryden Road
Dayton, Ohio 45439

Abstract

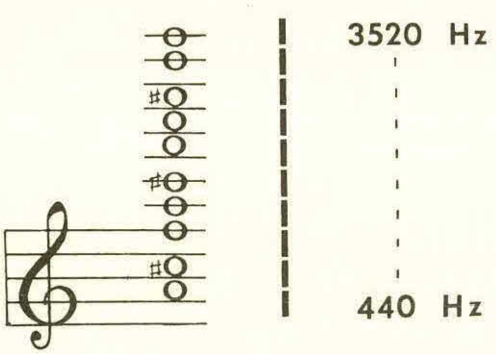
The Stereotoner is a new aural direct translation reading aid for the blind which is noteworthy for a stereophonic output code, a 10:1 zooming range, a switch which permits normal operation on reversed (light on dark) letters, a very small optical probe, and a compact, light weight control box which is suspended in front of the user's chest from an adjustable neckstrap. Maximum reading speeds have not been determined. Speeds of up to 50 words per minute have been reported. This paper discusses the engineering aspects of many of the Stereotoner features.

The work done and progress made toward manufacturing 100 Stereotoners also is described along with tentative descriptions of plans which will use 65 units in a realistic evaluation study.



How the Stereotoner Works

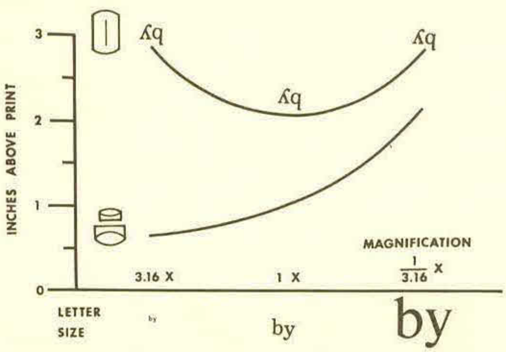
The Stereotoner probe is like a miniature camera with a built-in light except that the film is replaced by a vertical column of ten photocells which convert the light falling on them into electrical currents. When a photocell is darkened by part of a letter image, electronic circuits in the Control Box produce a musical tone. There are ten different tones, one for each photocell. Lower parts of letters produce lower pitched tones and upper parts produce higher pitched tones.



In stereophonic operation the higher tones appear to come from the user's right side and the lower tones from his left. In reading, the letters also seem spread from right to left. A capital letter "V", for example, would start with a high note from the right side followed by a series of descending notes moving toward the left. After the lowest point of the "V" is reached the notes move back up in pitch and back to the right side.

In like manner the user can identify the other printed or typed letters and the numerals by the pitch-time patterns of each assisted by the stereophonic effect. Shapes and symbols which have not been learned before can still be deciphered by working backward from the Stereotoner output.

In monaural operation, both earphones reproduce all tones at the same loudness so the stereophonic effect is not present. This mode of operation allows applying the complete set of tones to one earphone, to a monaural recorder, or to a telephone line.



The Stereotoner probe contains a unique variable magnification system which makes it possible to read more than a 10 to 1 range of letter sizes from 3/4" headlines to the smallest classified ads. A single knob on the side of the probe can be moved completely around the probe to adjust it for the letter size by automatically positioning the lens and the photocells to the correct distance from the paper. A lamp control knob atop the probe is used to adjust the light as required for the letter size, the paper whiteness, and the desired sensitivity to the print.

For reading italics the roller carriage attached to the probe base can be rotated. Moving the Reverse/Normal switch on the Control Box to Reverse makes the Stereotoner respond to light letters on a dark background as if they were dark letters on a light background. Individual volume controls for each earphone permit balancing the left and right outputs so that middle parts of the letters appear located centrally.

Using the Stereotoner

In use, the Stereotoner Control Box is suspended by its neckstrap over the user's chest and the probe is held vertically between the thumb and finger tips of one hand. Rolling the probe along the line of print produces the musical tone patterns associated with each letter.

Although the Stereotoner probe is most often used freehand, a straight edge Tracking Aid is supplied with each Stereotoner. This Aid helps the beginner roll the probe in a straight line and move it from line to line without losing alignment. The Tracking Aid is stored in a leather case which can be snapped to the rear of the Control Box for carrying.

The Stereotoner is used mainly for relatively short reading tasks. Even at the 35-50 words per minute achieved by some users, reading a complete novel would take too long unless there was no faster method available. The Stereotoner is well suited for reading typewritten personal letters, memos, bills, and notices. Short school assignments, lesson plans, etc. may also be read.

The capability of the Stereotoner probe for reading very large letters is useful for headlines, titles, and labels. Many box and can labels use large reversed letters. The reversed mode of operation is also useful for identifying the denomination of paper money. The capability for reading extremely fine print is useful in dictionary and reference book usage.

The probe size and shape are helpful in checking typing without taking the paper out of the typewriter. Non-reading uses include checking whether clothes or other objects are light or dark, checking ones own signature, and locating sources of light.

Training and Evaluation

Although a well qualified and highly motivated individual might learn to read well without outside assistance, most users have utilized a series of tape recorded lessons from Hadley School for the Blind and a week or two of personal instruction at either one of three V.A. centers, Hines, Ill.; Palo Alto, Cal.; or West Haven, Conn. (limited to veterans and potential instructors) or at Hadley.

For more information, write to The Hadley School for the Blind, 700 Elm Street, Winnetka, Illinois 60093. A more advanced series of lessons is also offered.

During the coming year up to 65 Stereotoners will be used in evaluation studies planned by Dr. Weisgerber of American Institutes for Research in Palo Alto, CA. A tape recorded test will be used for local screening of potential Stereotoner users. Those who appear suited to use the device will receive 2 - 3 weeks of training at one of the Centers listed above. AIR has also prepared the necessary training materials. Using weekly postcard replys, telephone and personal interviews, AIR will follow the students' further progress after this initial training to evaluate the effectiveness of the training program and the usefulness of the Stereotoner for each individual in his activities. In addition to the Stereotoners purchased by the Veterans Administration and the National Academy of Sciences, a number of units are being sold to individuals and institutions at a cost of $1,020.00 plus shipping from Mauch Labs in Dayton, Ohio.

Stereotoner Specifications

Output Code

Ten musical tones (A, C#, F) with the nominal frequencies of 440, 554, 698, 880, 1109, 1397, 1760, 2218, 2795, and 3520 Hz. Choice of mono (equal amplitude) or stereophonic (produced by different amplitudes of tones to left and right ears) outputs.

Control Box

Adjustable neckstrap with rubber pad, lined storage compartments for earphones, three switches for on-off, stereo-mono, normal-reversed print, left and right channel volume controls, left and right accessory output jacks for optional extra earphones for teaching, demonstrating, and recording, charger input jack, and snap fastener for Tracking Aid Case. Size: 5.4"Wx4.4"Hx1.9"D not incl. neckstrap. Weight: 19.5 oz., attaching the Tracking Aid adds 3 oz.

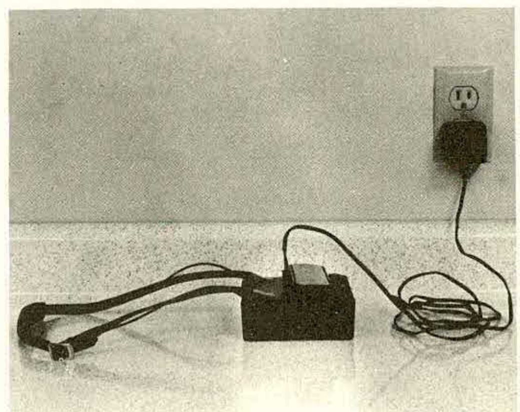
Power Source

User replaceable sealed NiCad rechargeable battery, 6.2 volt, 550 ma hr., approx. D size, 5 oz., 6-10 hrs. operation depending upon letter size, can be recharged in place hundreds of times. Power Consumption: 50-80 ma @ 6.2v (0.3-0.5 watt); for pica type consumption is approximately 0.4 watt. (About 8 hours)

Electronics: Standard, readily available solid state components assembled on three etched circuit boards; 41 transistors, 4 integrated circuits (containing a total of 80 transistors), 15 diodes, 101 resistors and 28 capacitors. Operating temperature: 32°F-115°F. Storage temperature: 0°F-140°F.

Charger Furnished

Plug-in, wall outlet mounted type with 6 foot low voltage cable with plug to fit jack in back of Stereotoner, 115 volts 60 Hz input, 50 ma dc output to battery, charges completely discharged battery overnight (14 hrs.) - 220 volt 50/60 Hz model may be substituted.



Stereotoner Optical Probe/Camera

10:1 zoom range from a magnification of 3.16 times to a reduction of 3.16 times covering a field of view from .075 to .750 inch and letter sizes from 6 to about 72 points, a single knob adjusts size and locks in position. 10 area photoconductors (CdSe) each .006 inch wide and .021 inch high in a column about .250 inch high sealed in a dry atmosphere. Four element lens 13 mm focal length, f 1.9 aperture. Single miniature lamp rated at 5 volts - 60 ma. and 3,000 hours life, usage at pica type size gives an expected life 20 times longer.

The work reported above is sponsored by the Prosthetic and Sensory Aids Service, Department of Medicine and Surgery, Veterans Administration.

THE NEW C-5 LASER CANE FOR THE BLIND

by

J. Malvern Benjamin, Jr.
Bionic Instruments, Inc.
221 Rock Hill Road
Bala Cynwyd, PA 19004

Introduction

In 1965 the first model of a laser cane for the blind was produced for the Veterans Administration (VA)* by Bionic Instruments, Inc. Since then, development has proceeded through three models (1) to the present C-5 Cane, the final "production" model to be described here.

During the past 35 years, many other efforts have been made to develop useful electronic mobility aids for the blind; perhaps twenty devices have been built and tested (2). In the last three years, three of these devices have shown sufficient promise to warrant detailed examination.

The simplest of these is the "Pathsounder" developed by Lindsay Russell (3). It consists of a camera-sized box, carried on the chest with a neck strap, which probes 6' (2 m) ahead with an ultrasonic beam. An obstacle detected between 30" and 72" (75 and 180 cm) from the device produces a low-pitched tone from small speakers mounted in the neck strap, while a high-pitched tone is emitted if the object is closer. This device thus protects its wearer against obstacles overhanging at chest to head heights that would not be detected by a cane or a dog. Since the device serves this single purpose, it must be used with a cane or a dog, but its use is easily learned and it is quite reliable.

Much more complex are the Kay Ultrasonic Spectacles (4, 5), a binaural device which, after suitable training, allows the wearer to locate objects in space. The sounds emitted into each ear blend, so that to an experienced user the objects themselves sometimes seem to be emitting the sounds. However, since the spectacles cannot detect dropoffs, they must be used in conjunction with a cane or a dog.

Intermediate in complexity, and consequently in training time required, stands the Laser Cane.

These devices, as well as several other sensory aides, are discussed in an excellent survey article on the subject (6).

* We appreciatively acknowledge the unfailing support of this project by the Veterans Administration through long and often difficult years. It has meant much to all of us who have worked on the project.

General Operation

As discussed above, the purpose of the Laser Cane is to add to the environment-probing ability of the conventional long cane (or "Typhlocane"). It provides three important pieces of information to the traveler, thus reducing his tension and making his progress more graceful.

Figure 1. Laser Cane in Use.

The Cane, shown in Fig. 1, emits pulses of infrared light, which, if reflected from an object in front of it, are detected by photodiodes placed behind receiving lenses. The angle made by the diffuse reflected ray passing through a receiving lens is an indication of the distance to the object detected. This principle of optical triangulation, first used by Cranberg (7) for that purpose in 1943, was chosen in preference to an ultrasonic approach because of its simplicity and small beamwidth.

The Cane's DOWN channel emits a 200-Hz tone to notify the user of any dropoff larger that 6" (15 cm) which appears approximately two paces (6' or 1.8 m) in front of the traveler. Some of the most serious of these would be down-going flights of stairs, the edges of train platforms, and open manholes and cellar-ways, but down-curbs, though more shallow, can usually be detected.

The FORWARD beam, about 2' (60 cm) above the ground, can range out to a maximum distance of 5' or 12' (1.5 or 3.5 m) from the cane tip to a light-colored target. The range is set by a lever located above the laser housing. Any obstacle detected within the selected range will actuate a stimulator that contacts the index finger when the cane is carried in the usual "long cane" manner. In addition, a 1600-Hz tone may also be switched on, if desired.

The UP beam will detect obstacles at head height appearing 1 1/2' to 2' (45 to 60 cm) in front of the cane tip. Obstacle detection is signaled by a high-pitched (2600-Hz) "beep."

Optics

The arrangement of the beams is shown in Fig. 2. Each of the dotted lines represents

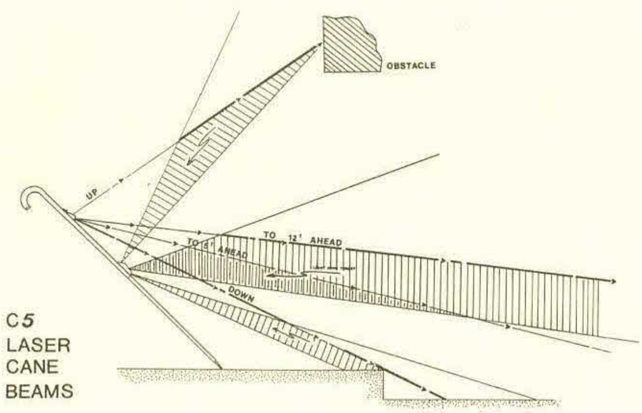


Figure 2. Laser Cane Beam Configuration.

a thin beam of pulsed laser light less than 1" (25 mm) in diameter at 10' (3 m). The cross-hatched fans depict the angular reception cones of the receiving lenses, which have to be much larger angularly than the laser beam in order to allow for variations in cane tilt, ground tilt, and range.

The DOWN beam is focused by a plastic lens 5 mm in diameter with a focal length of 13 mm; the other two beams are focused by lenses 6 mm in diameter with focal lengths of 7.6 mm. Emerging from these lenses, the beams, after redirection through front-surfaced mirrors, exit through two windows.

The receiving lenses (an 18-mm dia., 20-mm FL for the DOWN beam; two pairs of 12.7-mm dia., 17-mm FL (giving an EFL of 9.6 mm) for the other two beams) are spaced down the shaft from the upper nacelle so as to form a range base about 11" (28 cm) long. The angle that the reflected central ray makes with the central axis of the receiving lens thus indicates the range of the target.

The focal lengths of the receiving and transmitting lenses are in the same ratio as the photodiode diameter and the laser beam length. The receiving lens diameter is made as large as is practical in order to receive maximum light flux, while the transmitting lens diameter need only be large enough to subtend the solid angle required to capture all of the emitted light.

The UP channel optics is adjusted so that an obstacle approximately 18" (45 cm) in front of the cane tip and 6' (1.8 m) high will just be detected. Obstacles above this height will not be detected; so one will not get a false alarm when passing through a doorway. A twig the diameter of a lead pencil or a clothesline will usually be just detectable. As can be seen from the drawing, the detection area (that is, the line of laser light that is passing through the cross-hatched cone of the receiving lens) moves downward as one approaches the cane tip; thus the head is protected at the greatest distance, followed by the shoulders down to the waist. Detection range could, of course, have been increased, but it was determined in use that 18" (45 cm) was the best compromise between being warned too early, on the one hand, and not having enough time to stop at all, on the other. This channel will be actuated not only by overhead objects, but, of course, by objects which start at the ground and continue upward to head level. A wall, therefore, will first produce a FORWARD-channel signal as one walks toward it, then, within 18" (45 cm) of the wall, the UP-channel signal will also be activated.

Ranging on the FORWARD channel is done, as can be seen from the diagram, by shifting the angle of projection of the FORWARD-channel laser. This is accomplished by rotating a very small mirror through 3.5 degrees. The movement of the range switch mounted above the laser windows is transmitted to the mirror by a short, thin wire. By using mirrors, the laser assembly is compacted;

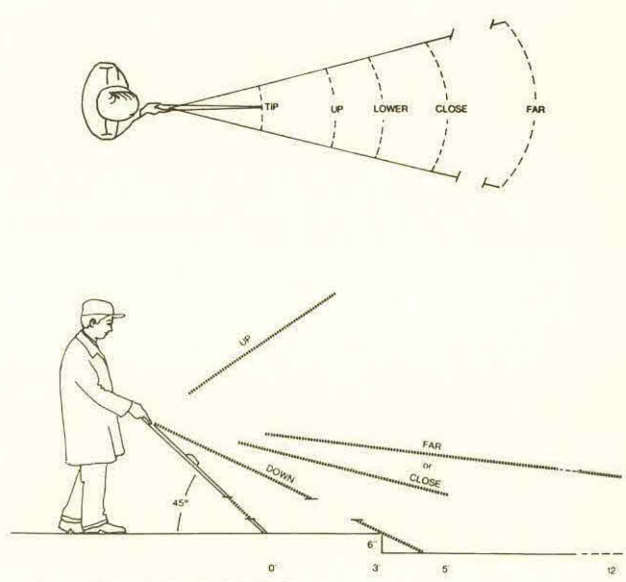


Figure 3. Zones of Protection of the Laser Cane.

the FORWARD and DOWN laser beams emerge from the lower window, while the UP beam uses the upper window.

For greater clarity, these sensitive zones of the cane are shown in the lower portion of Fig. 3. This representation shows that the entire length of the body is protected by one or another of the three beams, except for a small portion near the ground which is protected by the cane itself.

The upper portion of the figure is a top view, which illustrates more clearly the range at which the user is protected by each beam. Arcs are produced by the conventional "long cane" scanning technique of the user as he sweeps from side to side with each step.

The method by which the DOWN system detects a curb can be understood by studying Fig. 2. It will be noted that the laser beam strikes the ground at a more acute angle than the field of view of the receiving optics can accommodate because the laser assembly is mounted above the receiving lens assembly. The curb can thus be said to "cast a shadow" for the receive optics while the laser looks into this shadow. There is thus a momentary interruption in the otherwise steady stream of light pulses being received from the ground by the DOWN channel receiving system. The logic is arranged so that the system remains silent unless a pulse is missed, at which point a brief, low-pitched tone is emitted. For a walking speed of 4 mph (6.5 kph) or less, a repetition rate of 40 pulses per second is sufficient to assure at least one returned pulse for a 6" (15 cm) curb.

For this system to operate, it is quite necessary that both transmit and receive optics be quite sharp, as any significant amount of stray light will obliterate the "shadow" cast by the curb. At the same time, the dynamic range for the lower channel is great enough not to pick up too much stray light from a white concrete surface, while at the same time it is able to reflect sufficient light to keep the system from responding to a wet, shiny black asphalt surface. This range is approximately 300:1; thus the optical requirements for this portion of the system are severe. The present cane is capable of detecting a 6" (15 cm) curb under these conditions, but will give a false alarm on shiny black vinyl tile, polished black marble, or clean black asphalt coated with oil. If it were adjusted to miss these surfaces, it would then respond only to steps of 9" (23 cm) or deeper.

The lasers used in this system are Laser Diode Laboratory LD-23's with a typical peak power of approximately 12 W. However, only 2.5 W is radiated by each of the UP and FORWARD channels and 4 W by the DOWN channel. The laser safety situation has been studied by five separate laboratories; none has concluded that there is any danger (8).

PIN silicon photodiodes manufactured by United Detector Technology have been found to have the best signal-to-noise ratio and linearity over the greatest dynamic range, and so are used in this system. It is quite important that the gain in the photodiodes not change with ambient light and that the noise level not increase by much more than theory would predict in order that the system can have the highest signal-to-noise ratio practicable. This, in turn, is necessary in order to reduce size and weight.

Wratten 83C infrared filter material placed in front of the photodiodes further increases the signal-to-noise ratio by reducing the photodiode noise generated by high ambient light intensity.

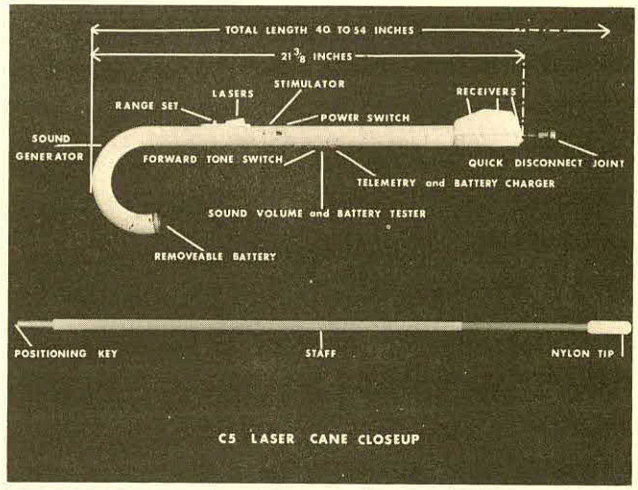


Figure 4. Close-up of the Laser Cane.

Mechanical Embodiment

Fig. 4 shows a labelled close-up of the C-5 Cane. The thin shaft is, of course, the lower section, which contains no electronics. Its tip is made of solid nylon and screws into the bottom of the staff for easy replacement as it wears. Tips of other materials have been explored, but they have the wrong frictional "feel" on different types of material, are too heavy, wear too fast, or do not transmit "vibration" in a manner familiar to the blind mechanical cane user. The nylon tip thus chosen for this cane is the same material that is conventionally used in the "Typhlo" long cane used by many blind people. Alternatively, a metal furniture glide tip can be provided for those who desire it. The staff can be supplied in either aluminum or fiberglass.

It is desirable that the shaft should plug into the upper electronic section of the cane with considerable ease so that it may be connected and disconnected rapidly, but it is also necessary that the joint wear well and that there be very little play. To achieve this, the joint is made with "O"-rings that press lightly against an aluminum ferrule inserted in the upper end of the shaft.

Cane length should be adjustable in approximately 1" (25 cm) increments over a range of 12" (30 cm) from 42" to 54" (105 to 135 cm). This is achieved by cutting the shaft to individual measure and then inserting the metal ferrule.

Just above the joint in the upper section appears a swelling containing the three lenses of the receiving optics. Halfway up and to the rear as the cane is being carried is a small panel containing several controls. The lowest is a jack used for charging the battery and also for coupling into the auditory output signals for telemetering purposes. This telemetering capability is quite useful during the training period, when the instructor may wish to trail the student by approximately one-half block. Output from this jack can also be fed to a hearing aid, if required.

Directly above this jack is the sound output volume control. Turned to one end, it operates a switch that turns on a tone whose volume indicates the amount of charge left in the battery. When the battery check tone can no longer be heard, approximately 30 minutes of use are left before recharge is necessary.

Directly above the volume control is a switch that turns on the tone associated with the FORWARD channel. The tactile stimulator is always in operation when the FORWARD channel detects an obstacle, but the tone is optional.

Continuing upward along the cane toward the crook, one next comes to the on-off switch, mounted in the right-hand side just under the position occupied by the right index finger of the user when he is carrying the cane using conventional "long cane" technique.

After operating the switch to turn the cane on, the index finger is returned to a resting position in which the ball of the finger resides in a small depression in the cane wall. At the center of that depression is a hole through which the tactile stimulator tickles the finger to warn of a forward obstacle.

Near the top of the cane, another forward-pointing swelling houses the lasers and transmitting optics. It contains a small switch that selects the FORWARD range. Pushing the switch down (that is, away from the user), sets the maximum range at 12' (3.5 m). Pulling it up toward the user reduces the range to 5' (1.5 m). Thus, indoors in cluttered surroundings or outside in a crowd, the range can be shortened enough to keep the obstacle detector from alarming the user unnecessarily.

The space inside the cane between the upper and lower optical housings is occupied by printed circuit boards containing most of the electronics needed to run the cane, while above the laser assembly is the laser pulse drive circuit.

Proceeding around the curve of the crook, a grillwork is located in the forward section of the cane. Beneath this grillwork are two tiny electromagnetic speakers which beam the sound upward to the user's ear. A special material with controlled microporosity allows virtually unattenuated sound passage while stopping water from penetrating into the cane.

The remainder of the crook houses a 6-volt 225-milliampere-hour NiCd rechargeable battery. This battery should last approximately three hours between charges and requires 12 hours to recharge. Its lifetime is one to two years, and it is easily replaceable with a new or freshly charged battery, when required.

The upper section of the cane is 1" (25 mm) in diameter. The whole cane weighs just over one pound (453 gm), as compared to 6 to 8 ounces (170 to 230 gm) for a conventional mechanical cane. The area occupied by the hand has been knurled to provide a rough grasp that will not slip if the hand becomes sweaty. A conventional smooth white enamel finish identifies the usual blind traveler's "white cane."

Electronics

The block diagram, Fig. 5, shows the electronic arrangement. Integrated circuits are used for the logic, but the rest of the system was built with discrete circuitry to save power and to save money in limited quantity production. When the total system is in operation with both speakers at full volume and the stimulator running, approximately 460 mW are consumed, of which the lasers and associated drive circuitry use 60 mW, the speakers and the stimulator use 350 mW, and the rest of the electronics uses 50 mW. Since so much power is required to actuate the output, the time between charges will be greatly dependent upon the amount of time that the outputs are turned on.

It will be noted in the block diagram that the laser drivers fire alternately instead of simultaneously. The UP and FORWARD-channel laser are connected in series, and so are driven simultaneously, while the DOWN channel is turned on only when the other two channels are off. This prevents the optical cross-talk that would be caused by light scattered from the UP or FORWARD channels being received by the lower channel; and it also, of course, reduces electrical cross-talk.

The lasers are pulsed at a 40-Hz rate for 200 nanoseconds. The receiving amplifiers are all gated so that they are turned off except when a signal can be accepted. This reduces the likelihood of interference from bright flashes of reflected sunlight from, say, an automobile windshield, and optical or electrical interference from fluorescent lights. Incidentally, two Laser Canes pointed at each other will respond only very occasionally!

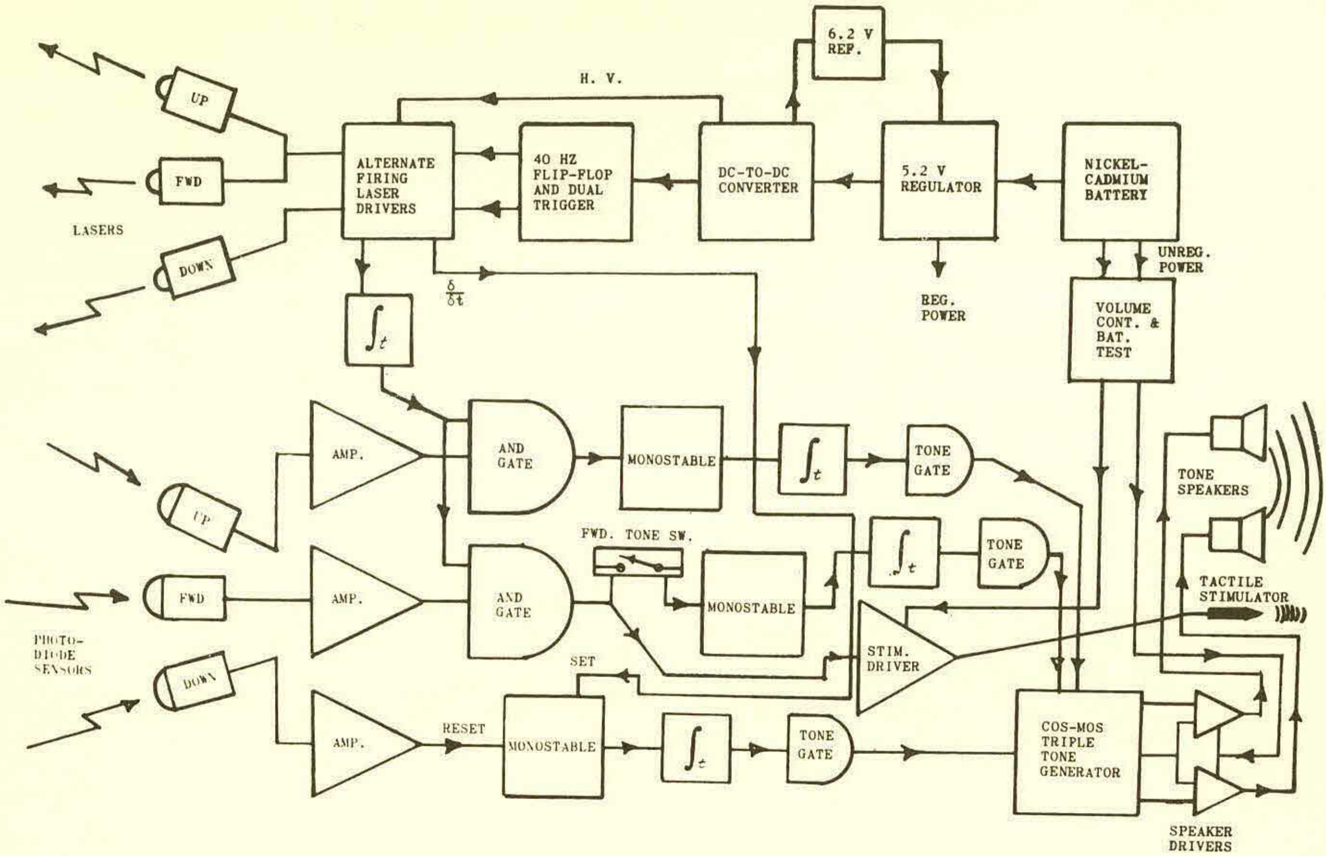


Figure 5. Electronic Block Diagram of the Laser Cane.

The amplifier bandwidth is approximately 5 kHz to 2 MHz to maximize signal-to-noise ratio.

Input signal level from a vertical white target at 5' (1.5 m) is 90 mV peak for the 0.2 μsec pulse. The poorest target material for which the system is required to work is clean black asphalt covered with a thin layer of water. Signal return from this material, viewed at a 50° angle, is only 0.5% of that for a vertical white target.

Outputs

The three cane output tones are anharmonically related to avoid blending. They are: UP, 2600 Hz; FORWARD, 1600 Hz, and DOWN, 200 Hz. To further aid in discriminating these tones, the waveform of the DOWN channel is tailored to produce a "rasping" sound, as compared to "pure" tones for the UP and FORWARD channels.

It was found that small sonic transducers, either piezo-electric or electromagnetic, are quite non-linear in frequency response and, of course, have poor low-frequency response also. Thus, if two or three signals are transmitted at once, intermodulation distortion occurs, producing entirely different tones. The user then must memorize not only the three solo tones for the three channels, but also their combination tones. In the Model C-5 this has been overcome by using two miniature electromagnetic speakers. One speaker transmits the high-pitched UP tone; the other transmits the middle-pitched FORWARD tone; and both transmit the DOWN tone. It is relatively rare that either the FORWARD or UP tones are sounded at the same time as the DOWN channel, while the UP and FORWARD channels are sounded together fairly frequently (as when one walks too close to a wall). Furthermore, by running both speakers at the same time, the poor low-frequency response inherent in a small electromechanical device can be compensated for in the DOWN channel. In addition, the acoustic chamber in which the two speakers are lodged was designed to resonate at this lower frequency.

The tactile stimulator is always connected into the circuit; so even if the volume has been turned down too low to hear or the FORWARD-channel audio tone has been turned off, the user will feel a tactile response to a forward obstacle. If his finger is not covering the stimulator hole, he will hear the tactile stimulator, which under those conditions will not be silent.

Present Project Status

Eight veterans have been using an earlier model (the C-4) of the Cane for one and one-half years with apparent satisfaction. Consequently, the Veterans Administration placed an order for forty more canes, to be finished by the end of 1973 for distribution and further study. The C-5 Cane described above has just been completed in prototype form, and forty copies of it will be produced. It is expected that in 1974, when the first production run of the cane can be made for general distribution, its price will be between $1,500 and $2,000. This compares with a price of $3,000 to $4,500 for a trained dog; and, like the dog, the cane will, of course, have to be subsidized.

References

1. Benjamin, J.M., Jr.: "A Review of the Veterans Administration Blind Guidance Device Project," Bull. Prosthetics Research, BPR 10-9, Spring, 1968, pp. 63-90.

2. Zahl, Paul A.: Blindness. Princeton, Princeton University Press, 1950.

3. Nye, P.W., and Bliss, J.C.: "Sensory Aids for the Blind: A Challenging Problem with Lessons for the Future," Proc. IEEE, Vol. 58, No. 12,. Dec., 1970, p. 1892.

4. Ibid., p. 1891.

5. Kay, L.: "Ultrasonic Mobility Aids for the Blind," Proc. Rotterdam Mobility Research Conf., May, 1965, pp. 9-16.

6. Nye and Bliss, loc. cit., pp 1878-1898.

7. Cranberg, Lawrence: "Sensory Aid for the Blind," Electronics, March, 1946, p. 116.

8. References available on request.

STUDIES IN GAIT IMPROVEMENT BY FUNCTIONAL STIMULATION OF THE LATERAL POPLITIAL NERVE

by

J. Vredenbregt
Institute for Perception Research
Findhoven - Netherlands

As the foot lifting muscles do not contract, the hemiplegic patient assumes a typical gait. To overcome this disability a simple electronic nerve-muscle stimulator has been developed. The start of the stimulation, as well as its duration, is controlled pneumatically by a rubber insole.

Because the moments at which the unaffected foot touches and leaves the ground correspond closely with those at which the affected foot must start and stop its swing phase, the insole is placed in the shoe under the ball of the non-affected foot.

A few years experience has demonstrated that a number of patients who used the appliance for a comparatively short period, have acquired a gait which can be qualified as almost normal to considerably improved. In a non-negligible number of patients the improvement remained even without further use of the appliance.

Reflex studies in some patients with this gait improvement show also an improvement of the recovery cycle of the Hoffmann reflex. This effect is probably caused by an increase in neural inhibition due to the functional stimulation.

Typical hemiplegic gait involves circumduction of the leg and dragging of the toe, due to weakness of the dorsiflexors. A patient with this disability (sometimes called drop-foot) is in danger of stumbling over his own foot. Therefore, when walking he tends to remain standing on the sound leg until the affected one is safely and securely on the ground. This instability during walking necessitates the use of a stick, but even with a stick he moves only with difficulty.

It has been found that this disability can often be largely overcome if an electric stimulus is applied at the appropriate instant to the nerve that causes the muscles involved to contract. A small portable electronic muscle stimulator has been developed and tested by the Institute for Perception Research in cooperation with the Rehabilitation Centre at Eindhoven, Netherlands.

Our main problem was to synchronize at the correct point of time the periodically supplied stimuli with the swing and stand phases of the affected leg during the patient's walking. In earlier investigations in this field, experiments were carried out with a mechanical switch under the heel of the affected foot (Liberson et al., 1961)[1]. It is our experience that the disadvantage of this arrangment was that the stimuli were switched on and off at times which came too soon with respect to the desired cycle of lifting and lowering the foot. Extensive gait studies taught us that in normal gait the pressure between the foot and the ground is almost reduced to zero when the ball of the foot leaves the ground. At that moment the body weight has been taken over by the other leg. Hence, a better place for the switch was under the ball of the affected foot, but owing to the vulnerability and rapid wear of the mechanical switch, which was placed outside the shoe (and since this only works properly on hard level ground), the solution was not satisfactory.

We have now obtained very good results with a sort of pneumatic switch by which the stimulation is operated. The switch itself is operated by the air pressure of a small air chamber in a rubber insole. The latter is normally placed in the shoe of the unaffected foot, because this foot shows a normal or almost normal pattern during walking. The insole fits in every shoe without any change of the shoe itself. The air chamber is situated under the ball of the foot; the instant at which the ball touches and leaves the ground or, to be more exact, the instant at which the weight of the body is placed on the insole and removed again, coincide with the starting and stopping of the swing of the affected foot. The instants at which the muscles of this foot contract and relax are thus controlled from the non-affected foot.

Figure 1 shows, in a schematic view, the appliance with the switch. The weight is 110 grams. The generator, i.e. the stimulator proper, is carried on a belt around the patient's waist or in a trouser-pocket. It supplies a rectangular pulse of a repetitive rate of 50Hz, 0.6 msec duration and a maximum amplitude of 100V (rated current 6mA). For reasons of convenience, only the amplitude of the pulse can be varied by means of a potentiometer, not the shape of the pulse. An elastic bandage immediately under the knee of the affected leg carries the two electrodes which are energized by the pulses. The positive (passive) electrode is made of electrical conducting rubber and is in direct contact with the skin. The other (active) electrode, to which a maximum of 100V can be applied, is made of silicon rubber with a core of stainless steel and makes contact with the skin through a damp sponge in order to provide an almost equal potential over the electrode area and to avoid high contact resistance.

The appliance is semi-automated by using a crystal as a pneumatic switch. This has the advantage that the stimulation starts automatically when the patient begins to walk and stops when he stops walking. Moreover, the adjustment of the insole regarding the pressure in the pneumatic system is not critical, which also contributes to an easy and convenient use. Even when triggering by the heel contact is preferable, which may occur in a few cases, the insole can be used under the heel.

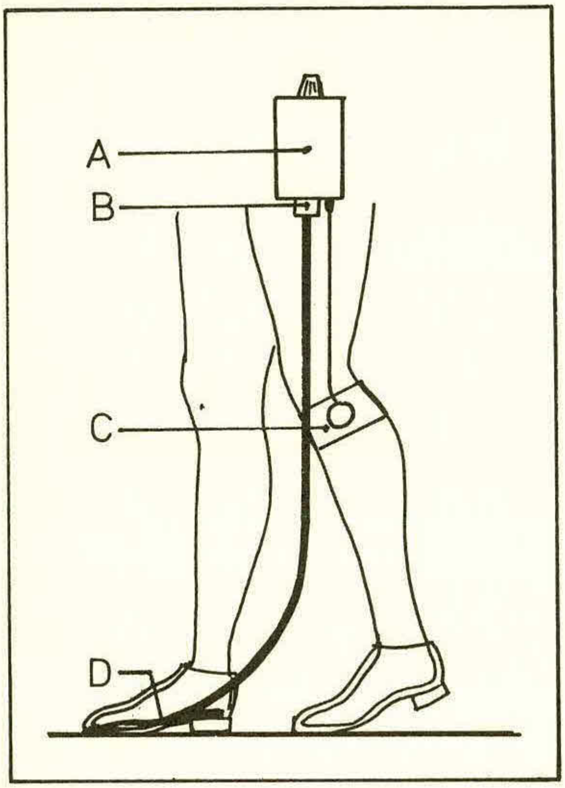


Fig. 1. Schematic view of the nerve-muscle stimulator.
A-generator, B-pneumatic switch,
C-stimulation electrodes,
D-insole.

The system has proven to be very reliable and the switch operates on any kind of floor cover or pavement without being damaged. This is a very important point, since it gives the patient confidence in the appliance.

The appliance has been used at various rehabilitation centres for a number of years. As a result, many patients have developed a more or less normal gait. An important contributory factor was found to be that the affected foot is better directed when the foot is placed on the ground before the stimulation is switched off and, therefore, stands more securely in the standing phase.

It is an interesting point to note that after some time a rather great number of patients, using the appliance, retain improvement in their walk even without further use of the appliance. As an example, 45 of 60 patients over 40 years of age who underwent the treatment for 3 to 6 weeks, have recovered completely or showed considerable improvement in their walking. Similar results were also reported by Liberson et al., 1961[2] and Vodonik et al., 1965[3]. The rest of the patients responded well while walking with the appliance, but showed no lasting improvement. A representative number of the last group have been using the appliance constantly for more than two years and are walking with confidence.

The fact that a non-negligible number of patients showed more or less complete improvement of their gait is yet one of our subjects of interest. To gain some insight into this phenomenon, spinal reflexes of the Achilles tendon and reflexes of the M. gastrocnemicus (Hoffman reflex) are being studied in cooperation with reflex specialists. Evoking spinal reflexes has the advantage that one is stimulating a comparatively well-known structure in the nervous system. As an example, figure 2 shows the recovery cycle of the Hoffman reflex of the legs of a right-sided hemiparetic patient, as found by Brunia (1971)[4]. A more rapid recovery of the Hoffman reflex is demonstrated in the affected leg, which was also reported by Diamantopoulos et al. (1966)[5], Magladery, (1952)[6] and Takamori, (1967)[7]. It appeared that a higher facilitation level is reached in the hemiparetic limb. This concerns both the Hoffmann and tendon reflexes.

Recent preliminary reflex studies in patients, demonstrating lasting gait improvement, also indicate an improvement of the recovery cycle. This indication could be derived from the fact that the recovery cycle of the reflex of the affected leg before and after using the appliance differs less from that of the non-affected leg before the appliance was used. The preliminary conclusion is that a decrease of the facilitation or an increase of the inhibition is initiated by applying functional stimulation.

The appliance is also being used on children with a drop-foot disability due to a brain defect contracted before or during birth. These children were hardly able

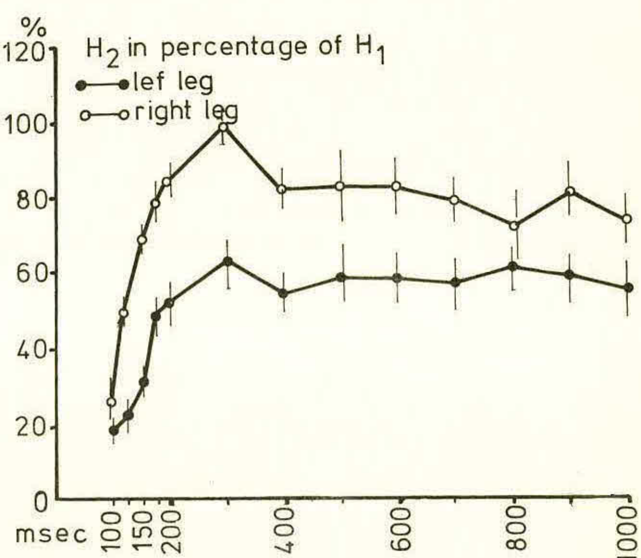


Fig. 2. Recovery cycle of the Hoffmann reflex (H) of a patient with right-sided hemiparesis. Every point of the curve is the mean of 18 responses (Brunia, 1971).

to learn to walk normally. Some of them were paralyzed at only one side and have been under treatment for several months and show an improvement in gait.

The results obtained with the appliance have demonstrated its helpful effect in a number of cases. However, it should be mentioned that functional stimulation will always have its restrictions.

References

1. Liberson, W. T., Holmquest, H. J., Soct, D. and Dow, M., "Functional Electro-therapy Stimulation of the Peroneal Nerve Synchronized with the Swing Phase of the Gait of Hemiplegic Patients," 1961, Asch. phys. Med. 42: 101-105.

2. Vodonik, L., Long, Co., "Myo-Electro Control of Paralyzed Muscles," 1965, Transactions IEEE, Vol. BME-12, 3-4: 169-172.

3. Vodonik, L., "Funktionelle Elektro-Stimulation/Orthotik," 1965, Medizinal Marktacta Medicotechnica, Vol. 8: 371-372.

4. Brunia, C.H.M., "Experiences with Spinal Reflexes in Research of the Human Motor System," 1971, Medicine and Sport, Vol. 6: Biomechanics II, 29-47, (Karger, Basel).

5. Diamantopoulos, E. and Zander Olsen, P., "Motoneuron Excitability in Patients with Abnormal Reflex Activity," Granit Muscular Afferents and Motor Control, 1966, Nobel Synposium I: 451-452, (Almqvist & Wiksell, Stockholm).

6. Magladery, J. W., Teasdall, R. D., Park, A. M., and Langhuth, H. W., "Electro-Physiological Studies of Reflex Activity in Patients with Lesions of the Nervous System," 1952, International Bulletin, Johns Hopkins Hospital: 219-244.

7. Takamori, M., "H-Reflex Study in Upper Motoneuron Diseases," 1967, Neurology 17: 32-40.

A MINIMUM ENERGY EXOSKELETON

by

Eugene V. Kosso
Electrical Engineering Department
University of Nevada
Reno, Nevada

Introduction

For the past few years work has been progressing at the Institute Mihailo Pupin in Belgrade under the direction of Prof. Rajko Tomovic on the design and construction of a powered exoskeleton for the rehabilitation of paraplegic and hemiplegic patients. The locomotion is under the control of a computer algorithm for walking and several other ambulatory functions. The general algorithmic control is described by Vukobratovic (1).

The major effort to date has been to stabilize the gait by imposing a motion on the upper torso. The device has been demonstrated with healthy and a few paralyzed subjects and while it has been demonstrated that a gait can be imposed on a subject which is stable for a few steps, it would appear that considerable work remains to be done before the device is accepted by the patients and their physicians as a natural aid to rehabilitation. The principal problems to be overcome are:

1. The general reluctance of patients to accept artificial aids. This reluctance is compounded by the slightest undependability of the device. If a patient has had a bad experience with an orthosis he likely will never use the device again, no matter what improvements are made. This reluctance is magnified in the case of an externally controlled device such as is described here.

2. The machine-man interface is critical. Since the locomotion is imposed externally on the patient the force must be transmitted to the limbs. Particular care must be exercised in the fitting at the force points to prevent dangerous bruising of the patient. Unfortunately, paralyzed persons seem to be particularly susceptible to bruising, and it is liable to have dangerous, lasting effect.

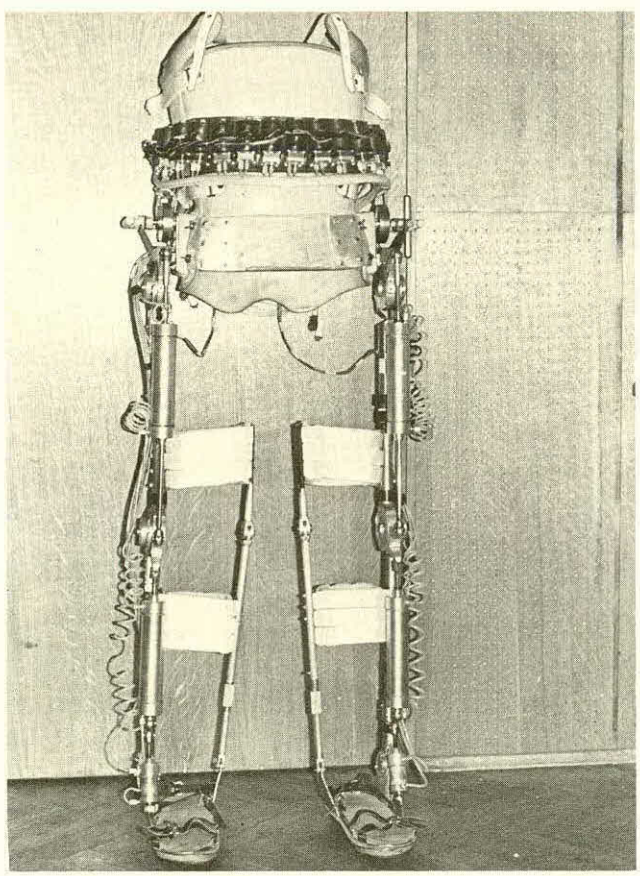


Figure 1.

3. The feasibility of the concept has been demonstrated, but much work remains to be done before the "walking machine" is an accepted orthosis. While an attempt has been made to study natural gaits and duplicate them via computer programs, the gait algorithms demonstrated to date appear to be far from natural. This leads to further reluctance of patients to accept the orthosis.

4. As is probably evident in Fig. 1, the orthosis is pneumatically powered, with the six cylinders operating the joints. They are controlled by the servo-valves around the upper body corset. This necessitates an external pressurized gas bottle with its attendent umbilical cord connected to the orthosis. This, of course, results in a limited range of ambulation.

Proposal

If the device is to be an autonomous system it will probably have to be redesigned to be electrically powered. To make this development possible it will be necessary to develop small, powerful electric motors to operate the joints. At the same time the energy consumption should be minimized if the energy source (batteries) is to be carried by the user. This necessitates the choice of proper, minimum energy, gait algorithms, and the conservation of energy by returning to the source some of the expended energy. It is this concept that is the subject of this report.

Procedure

If we examine the energy flow in the joints of the lower limbs during ambulation, such as can be derived from Peizer, et al., (2) for the kinetics of the knee, it can be noted that during a large part of the gait cycle the energy flow is negative, that is the muscles are actively used to "brake" the body momentum. In the body, as in the pneumatic system, this energy is not recovered, but leads to further expenditure from the energy source. However, it might be possible to recover and return to the source this energy in an electric system, much as an electric locomotive or elevator can return energy to the power system on a downhill trip, by using regeneration.

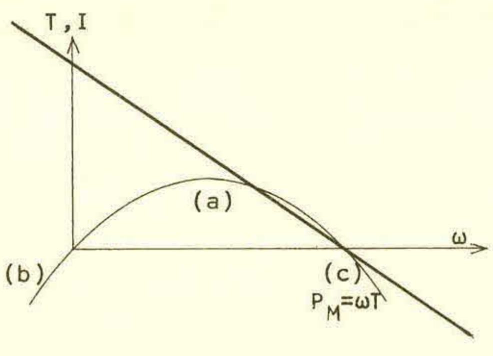


Figure 2. For constant-field servo-motor, normally operating on part (a) of the mechanical power curve.

Recovery of Negative Energy

Conventionally, electric servo-motors are either field-controlled, or more usually, armature-controlled with constant field current or permanent field magnets. However, this does not permit the reversal of electric energy flow, since, when the mechanical energy flow is reversed the result is an increase in internal electric losses so that instead of P_{elec} being reversed it is merely increased. This is illustrated in Fig. 2.

If the external load forces ω negative with $+T_{dev}$ the current increases further and the electric power source must now supply the increased I^2R losses plus the external mechanical load (part B of the curve).

In order to reverse the electric energy flow, it is necessary to operate on part (c) of the curve. In simple terms, in order to recharge the power pack, the generated voltage must be greater than the terminal voltage, with the same polarity. To achieve this the following scheme of field and armature polarity control is proposed.

Figure 3 indicates a method proposed for recovering the electric energy when the mechanical energy flow is reversed (mechanical power aiding or retarding the desired direction of motion). The logic of the operation is indicated in Fig. 3 and Table 1.

Referring to these diagrams, the basic (desired)

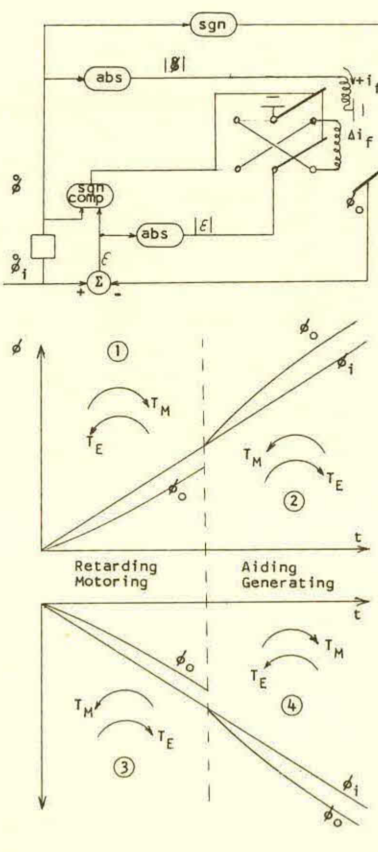


Table 1

	①	②	③	④
$\dot{\phi}_i$	+	+	-	-
T_M	-	+	+	-
v_a	+	+	-	-
i_a	+	-	-	+
T_E	+	-	-	+
Δi_f	-	+	-	+
ε	+	-	-	+

Figure 3.

angular velocity of the machine is set with i_f, which is always positive. Frrors of angular velocity are manifested in Δi_f. Depending on whether the load torque is aiding or retarding the desired direction will determine the direction of the increment of i_f, as indicated in the logic table. The direction of rotation is determined by the polarity of v_a. It will be observed that when the external load aids the desired direction of rotation $v_g > v_a$ so that the machine is generating, and electric energy flows to the armature supply. When the external load retards the desired direction of rotation $v_g < v_a$ so that the machine is motoring.

A block diagram for investigating the operation and stability of the system is given in Fig. 4. It is immediately noticed that, even if the flux-field current characteristic is linear, the system is fifth-order, as indicated by equations (1), describing a set of state-variables for the system, and nonlinear due to three multiplications of variables. This is undoubtedly a powerful reason for using only armature-controlled servo-motors, carefully linearized.

However, in the interest of engineers with applications requiring minimization of energy usage, such as for systems such as described herein, the system was analoged to determine the mode of operation and the possible stability or instability.

$$\left.\begin{aligned}
\mathring{x}_1 &= x_2 \\
\mathring{x}_2 &= -\frac{B}{J}x_2 + \frac{M}{J}x_3x_5 + \frac{M}{J}x_4x_5 - \frac{1}{J}T_L \\
\mathring{x}_3 &= -\frac{R_1}{L_1}x_3 + \frac{1}{L_1}\mathring{v}_i \\
\mathring{x}_4 &= -\frac{1}{L_2}x_1 - \frac{R_2}{L_2}x_4 + \frac{1}{L_2}v_i \\
\mathring{x}_5 &= -\frac{R_a}{L_a}x_5 + \frac{MR_1}{L_aL_1}x_1x_3 + \frac{M}{L_aL_2}x_1^2 \\
&\quad + \frac{MR_2}{L_aL_2}x_1x_4 - \frac{M}{L_a}x_2x_3 - \frac{M}{L_a}x_2x_4 + v_a
\end{aligned}\right\} \quad (1)$$

where:

$$\begin{aligned}
x_1 &= \phi_o \\
x_2 &= \mathring{\phi}_o \\
x_3 &= i_f \\
x_4 &= \Delta i_f \\
x_5 &= i_a
\end{aligned} \qquad (2)$$

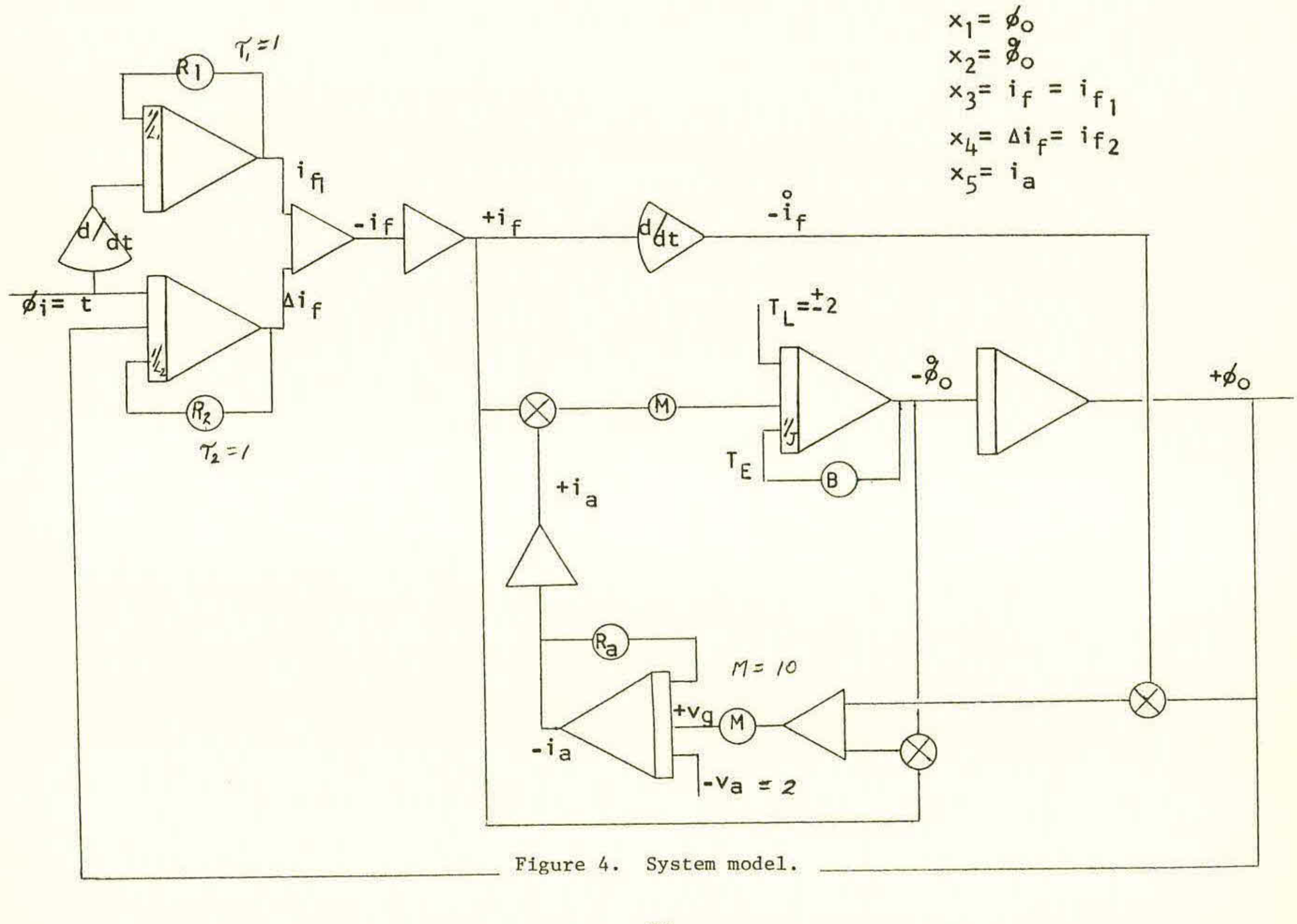


Figure 4. System model.

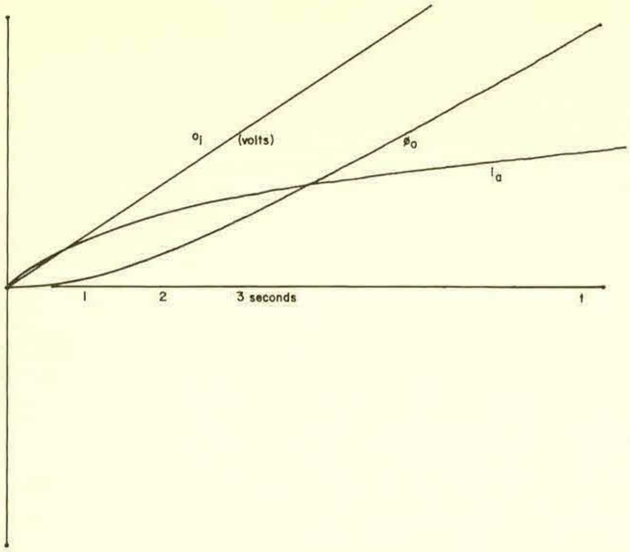


Figure 5. Mode 1.

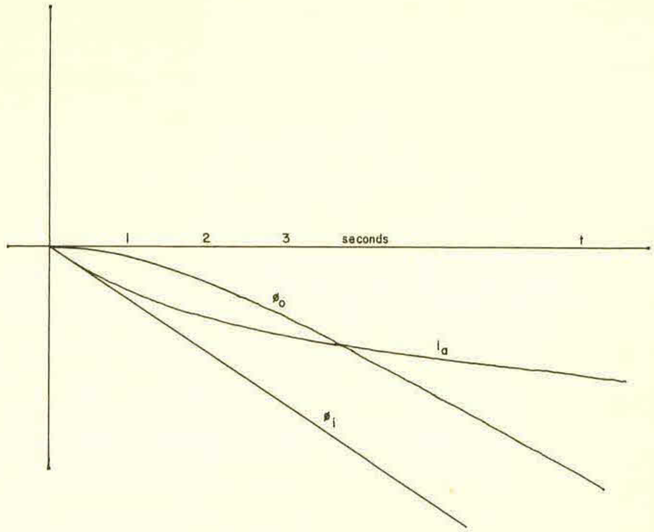


Figure 7. Mode 3.

Results

The results of an analog computer study are indicated by Figs. 5-8, on which are plotted the input angle; ϕ_i, the output angle; ϕ_o, and the armature current; i_a, for the four possible combinations of input angle and load torque. For these results, nominal parameter values were used, as indicated.

These results could be characterized as a positive "possibility" study. Since the armature current reversed for plots 6 and 8, the indication is that the power flow has been reversed for these cases. Also, at least for the values of parameters chosen, the system was stable.

At the present time, what might be characterized as a "feasibility" study is being carried out. An effort is being made to generate Lyapunov functions to determine the maximal region of state space in which the system model is stable and also the region of parameter space in which the system model is stable.

If the system model is stable over the range of variable and parameter values of interest, a physical system will be constructed and tested to determine if a significant energy economy is indeed obtained.

References

1. Vukobratovic, M., Frank, A. A., Juricic, D., "On the Stability of Biped Motion," IEEE Transactions on Bio-Medical Engineering, Vol. BME-17, No. 1, Jan 1970.

2. Peizer et al., "Human Locomotion," Bulletin of Prosthetic Research, Fall 1967.

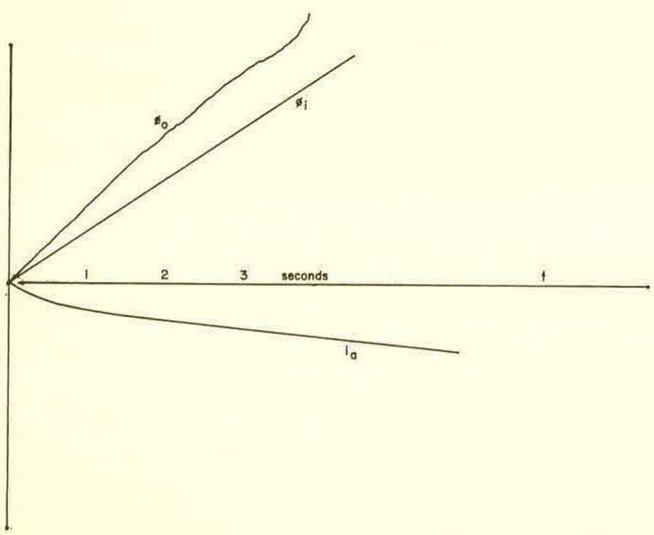


Figure 6. Mode 2.

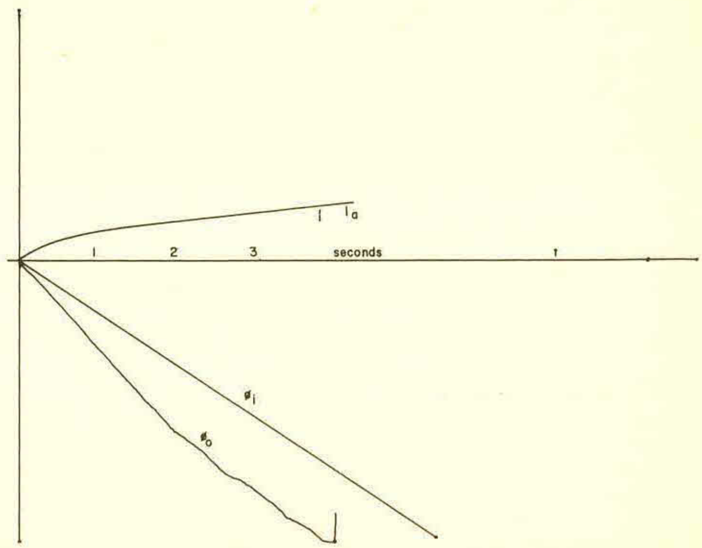


Figure 8. Mode 4.

STOCHASTIC ANALYSIS OF EMG SIGNALS FOR MULTI-FUNCTIONAL PROSTHESIS CONTROL PURPOSES

by

Daniel Graupe, William K. Cline and Ted K. Kaplon
Department of Electrical Engineering
Colorado State University
Fort Collins, Colorado

Summary. A stochastic analysis of EMG signals is described in terms of a time-series analysis. Following an analysis of the p.d.f. of the innovation of the EMG signals, an ARMA model for various recorded EMG signals is derived, and subsequently a K-B filter model is obtained. It is demonstrated that recognition of various limb funtions is facilitated via this analysis and the resultant identification, as is required for multi-functional EMG control of prostheses. Realization problems are discussed.

1. Introduction

A major problem in reliable multi-functional prosthesis control via EMG signals is that of the filtering of the EMG signals from noise and from other interacting EMG signals and of recognizing the EMG signals related to different limb functions. For adequate and reliable such filtering and recognition, a rigorous statistical EMG analysis is essential. However, the EMG analysis presently available in the literature is mostly restricted to ad hoc methods of spectral density and correlation evaluations [1]-[9]. Noting the recent progress in applying time-series analysis techniques to EEG analysis [10]-[12], a similar philosophy is presently being attempted to be applied to EMG analysis, noting that d.c. levels or amplitude analysis alone cannot facilitate separation of various interacting signals.

1.1 Hypotheses

The present approach to EMG analysis recognizes the non-stationary nonlinear nature of the EMG signals. Furthermore, it recognizes the practical constraints in any realization of filtering and identification or recognition, when employed in conjunction with a realistic prosthesis, these constraints being those of time duration available for processing and of cost and weight of computational hardware to be carried by the amputee (say, in his pocket). Certainly, these constraints must and will lead to compromises and to simplifying assumptions in any analysis that will detract from the rigor of the analysis. The present investigation is just aimed at checking if the compromises made may still yield adequate and reliable recognition, filtering and (subsequently) control.

The present analysis is based on the following hypothesis: (I) The EMG signals related to at least some different limb functions (not necessarily all functions) are detectable (by surface electrodes). (II) at least some parameters, that can be identified for the EMG time-series that are recorded, differ significantly in value or combination of values for different limb function. These two hypotheses imply that the information on at least some different limb functions is available and resolvable by identifying the time-series parameters. In other words, if we wish to resolve say, three different functions (functions a,b,c), then we assume that if we record the EMG signal at several locations (somewhere along the upper arm, shoulder or neck - for upper elbow prosthesis use), at some of these locations the time-series parameters have a certain stationarity feature such that at least some of the parameters of the various functions do not change in time so much that the various functions become indistinguishable. In other words, for a two function case (functions a,b), the range of time variation of the parameters for function a is such that the parameter space which the parameters of a may occupy at some time does not overlap that of function b (at least one parameter of a will not overlap or be overlapped by the corresponding parameter of b). The hypotheses above will all be tested, since if they prove to be incorrect, the present analysis will be useless.

We comment that the parameters considered above need not be similar for similar functions for different patients. To overcome this, complete offline identification (to set parameter ranges) must be made for each patient prior to connecting his prosthesis controller to the identification and filtering micro-computer that is to execute the recognition.

Obviously, since external or biological non-EMG noise (say noise due to ECG, to fluorescent lights, etc.) may have overlapping parameters, this should be filtered first.

2. Experiments

For the present analysis, EMG signals have been recorded at three different locations (L1,L2,L3) along the upper arm of a volunteer. (Later the same was repeated on a second volunteer.) The signals related to different limb functions: F1; complete rest (no arm bending), F2; load of one pound on hand during arm bending, F3; squeeze action (no arm bending) and F4; same as F2 but five pound load - See Tables of Section 4.

Functions F2, F3, and F4 were selected to examine if various loads for the same function may not be confused with another function, and to check how load change affects the hynamic parameters of the time-series.

We comment that although an amputee cannot perform the above functions (without a prosthesis) he can, or can be trained, to produce (above the amputation) muscular tensions and flexions related to the above functions.

This work was supported in part by the Veterans Administration, Department of Medicine and Surgery, Prosthetic and Sensory Aids Service, under contract V101(134)P-111.

The sampling rate of the EMG when fed to the computer was 2000 samples/sec which is well above the relevant upper frequency limit of interest in EMG, namely 1000c/s [7], the lowest frequencies containing useful information being about 30 c/s.

3. Analysis

Our analysis was concerned with steady state data. Obviously analysis of transient data is more appropriate if one wishes to actuate prosthesis movements by EMG signals. However, without any steady state analysis otherwise available, the authors felt that they should first analize the steady state and leave the transient analysis to a later stage after the steady state has been studied. Furthermore, it was hoped that identification may be complete after about 0.15 - 0.25 sec of steady state data, noting that the lowest frequency containing information is about 30 c/s [7], and that such a delay in actuating a prosthesis may be acceptable. We will return to this point in Section 5 below.

3.1 Test for Distribution

The constraints mentioned in Section 1.1 make it obvious that if a linear analysis can be even weakly justified, it should be used, since otherwise computation will be too lengthy and complex to enable practical prosthesis application. However, a linear model is the correct one only in the Gaussian case. Hence, and to avoid the complex testing of the joint distribution, our testing was as follows:

Let y_k denote the measured time-series ($k = 0,1,2,\ldots,$ denoting time interval). Noting [13], the whitening filter for y_k of any distribution is a linear pure autoregressive (AR) filter which is in theory of infinite order, but whose parameters converge for bounded y_k. Since this is our case, it was sufficient to use 15-20 AR terms to obtain a white residual as follows:

$$y_k = \sum_{i=1}^{s} \alpha_i y_{k-i} + w_k \qquad (1)$$

α_i being the AR parameter, s being the order of AR filter and w_k being the white noise residual. Eq. (1) can be written in operator form as

$$\alpha(B)y_k = w_k \qquad (2)$$

B being a delay operator; namely, $B^i y_k = y_{k-i}$. The parameters α_i where identifiable as in Chapter 12 of [13] using a least squares algorithm. Their identification can be shown to be consistent (convergent in probability) for the correct AR order, and to otherwise yield an error that can be bounded as required by appropriate selection of a [14], [15], [16].

Subsequently to the whitening, the marginal p.d.f. (probability density function) was checked, noting that if the p.d.f. turns out to be Gaussian, w_k is independent and no joint p.d.f. is required. Figures 1-a, 1-b indicate that w_k is in fact very close to Gaussian for various limb functions and rest situations. Although the p.d.f. is not exactly Gaussian (and was not expected to be), the constraints mentioned in Section 1.1 and earlier in Section 3.1, at least justify a linear approach to be adopted. (We comment that in an experiment that has been made, EMG reconstruction via the Gaussian assumption has yielded a prediction error of less than one percent.)

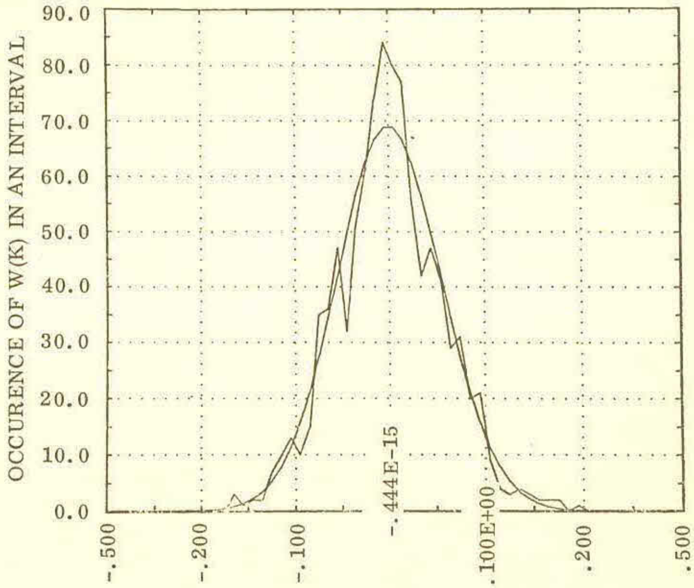


Figure 1a: p.d.f. of w_k at rest.

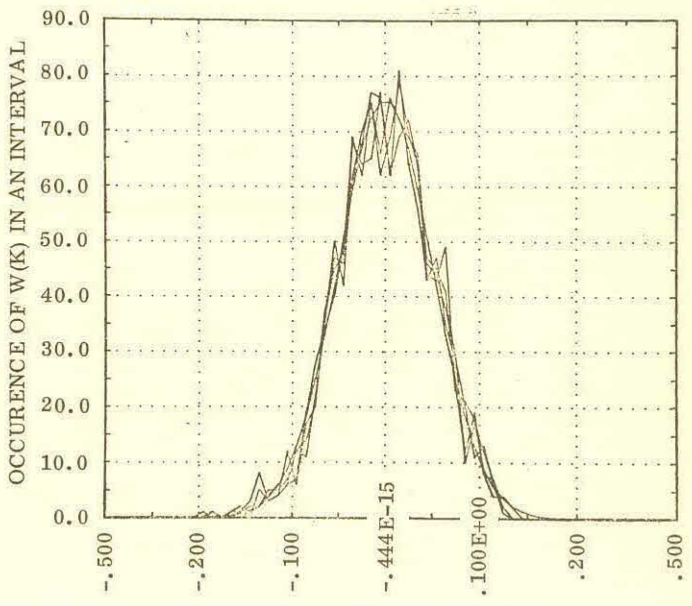


Figure 1b: p.d.f. of w_k for various limb functions.

3.2 Identification of Input/Output ARMA Models for the EMG Signals

Once the Gaussian assumption has been justified as a reasonable approximation, as was shown in Section 3.1, the way is open to identify linear EMG models. This model will be derived in terms of an ARMA (autoregressive-moving-average) formulation, which is transformable into a KB (Kalman-Bucy) filter form for further filtering of noise. We note that the KB filter is optimal in the Gaussian case and optimal linear otherwise. The identification of the ARMA model follows exactly the procedure outlined in Section 12.6 of [13] and in [15] and will not be repeated here. This procedure yields the following ARMA model:

$$\phi(B)y_k = \theta(B)w_k \tag{3}$$

w_k being the white noise (innovation) process of (1) and $\phi(B)$, $\theta(B)$ denoting the n'th order AR and the MA (moving-average) parameters of the ARMA model. Assuming that $\alpha(B)$ of (1),(2) are consistent (by [14] see Sect. 3.1 above) the model of (3) can be consistently identified noting [17] and theorem 2.3.3 of [18].

The procedure above is applied to all EMG recording locations such that for location μ ($\mu = 1,2,\ldots,r$) y_k of eqs. (1)-(3) becomes $y_k^{(\mu)}$, $\phi(B)$, $\theta(B)$, $\alpha(B)$ become $\phi_\mu(B)$, $\theta_\mu(B)$, $\alpha_\mu(B)$ respectively, and the $^\mu w_k$ above becomes $w_k^{(\mu)}$. The complete EMG model thus becomes a multi-output model, whose outputs are spatially distributed. The treatment of each output is, however, a single-input-single-output treatment. A multi-input-multi-output treatment, whose theory is outlined in [19] is certainly possible and will be more complete but it is far lengthier. It therefore is not recommended as long as the single-input-single-output applied to all $y^{(\mu)}$ facilitates adequate resolving of functions, and noting the constraints of Section 1.1. We whall return to this point in Section 5 below. For results and their discussion - Section 4.1

3.3 Derivation of Kalman-Bucy Filter Models for the EMG Signals

In Section 3.2 the identification of the ARMA model for EMG signal was discussed. However, y_k of eqs. (1)-(3) contains non-EMG measurement noise. This can be expressed as follows:

$$y_k = x_k + n_k \tag{4}$$

n_k being non-EMG measurement noise. Due to the nearly deterministic form of ECG, any ECG interference can be filtered out and is considered not to be white noise. This assumption is rather restrictive and can be relaxed. However, it is convenient (and very common) to filter some measurement noise. Consequently, and since the results available at this time are related to this assumption, we shall restrict the present discussion to n_k being white noise. Hence, substituting x_k from eq. (4) into (3) yields [20]:

$$\phi(B)x_k = \theta(B)u_k \; ; \quad E[u_k v_\ell] = 0 \; \forall \, k, \ell \tag{5}$$

u_k being white noise, $\theta(B)$ being of order $m \le n$, and

$$\sigma_w^2 \sum_{i=0}^{n-h} \psi_i \psi_{i+h} = \begin{cases} \sigma_n^2 \sum_{\alpha=1}^{m-h} \theta_\alpha \theta_{\alpha+h} + \sigma_v^2 \sum_{i=0}^{n-h} \psi_i \psi_{i+h} & (\forall\; 0 \le h \le m) \\ \sigma_v^2 \sum_{i=0}^{n-h} \psi_i \psi_{i+h} & (\forall\; m < h \le n) \end{cases} \tag{6}$$

θ_0 being 1. The solution of eq. (6) noting from Sect. 3.1 that w_k tends to a metrically transitive process for appropriate s to yield a consistent estimate of σ_w^2 [21],[22] gives consistent estimates of θ_i, σ_u^2, σ_v^2. Once these are available, transformation into the state space form required for K-B (Kalman-Bucy) filtering is straightforward [18], and all the required parameters of the K-B filter are available. Hence, linear-optimal (and in the Gaussian case, optimal) filtering is possible. Howver, for our EMG problem, the derivation of the ϕ, θ parameters as in Section 4 is our main purpose. For results and their discussion - see Section 4.2 below.

4. Computational Results for Recorded EMG Data

4.1 ARMA Models of y_k (eq.3)

EMG data for the experiments discussed in Section a (and for which the results of Section 3.1 are valid) have been employed to derive the models of eq. 2. The identified models obtained are tabulated in Table 1 - below for various assumed AR orders s. In these tables L1, L2,... denote locations 1,2,... and F1, F2,...denote functions 1,2,.... Whereas, $\hat{\phi}_i$, $\hat{\theta}_i$ denote the identified parameters of $\phi(B)$, and $\theta(B)$, where $\phi(B)y_k \equiv y_k + \phi_1 y_{k-1} + \ldots$,etc.

Table 1(a,b,c,d) indicates that already for small s (s=5) most models have settled to the neighborhood of the ARMA parameters and error variance (σ_w^2) obtained for high s. Hence, and noting that high s involve long computation time and long samples, the ARMA model for s=10 is adopted which hardly improves when s is further increased. In fact, improvement is evident especially for L2,F1 but the resulting parameter changes are far smaller than the changes for even high s between identifications obtained for L2,F1 at different times. Similarly, and examining different orders of ARMA models we can conclude that s=10 with an ARMA model where $\phi(B)$ and $\psi(B)$ are both of order 2 is both adequate and convenient. The effect of increasing s is demonstrated by observing AR models for different s. For this purpose we consider only L1,F1 (see Table 2), since the other location and functions yield similar behavior with respect to increasing s. We also note that the identification of Table 1 at location L1 is for AR and MA orders of n=2, whereas at L2 for F2,F3 it is for n=3. A discussion on order determination is given in [13],[15]. Consequently, and since in the present case, AR and MA orders of n=3 at L1 (and L2,F1) give hardly any reduction of variance, orders of n-2 were selected at that location. At L2 (for F2,F3) orders n=3 gave a considerable variance reduction, whereas n=4 gave no further improvement. Hence, n=3 was chosen there. Other criteria discussed in [13],[15] similarly point to the same choices of orders as above. From [15] and from that fact that eqs. (3), (4) must lead to eq. (5), it can be easily shown that $\phi(B)$ and $\psi(B)$ must be of the same order.

Table 1 (say 1-C) indicates that each of the functions F1,F2,F3 is distinguishable from the others by its parameter pattern in at least one location considered, and sometimes in more (since some of the parameters or of their combinations are sufficiently distinct from the respective parameters for other functions). This argument is further established if we consider the change in the parameters for the same function and location when using data recorded at different times, as shown in Table 3 for L1, s=10, where the ranges observed in four separate recordings are given (obviously, the number of

TABLE 1-a Model of Eq.(3); s = 3

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\sigma_w^2 \times 10^{-3}$
L1,F1	- .6750	- .2919		+ .1505	- .1159		4.452
L1,F2	-1.772	+ .8235		- .3937	+ .2071		1.250
L1,F3	-1.922	+ .9505		+ .0434	+ .0021		2.570
L2,F1	- .7686	+ .2009		+ .1592	+ .0170		5.436
L2,F2	- .6009	- .2764	- .1117	+ .3174	+ .0298	- .0210	6.839
L2,F3	-1.676	+1.210	- .3200	+ .0357	+ .0217	+ .0083	7.365

TABLE 1-b Model of Eq.(3); s = 5

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\sigma_w^2 \times 10^{-3}$
L1,F1	- .5296	- .4325		+ .2991	- .1313		4.452
L1,F2	-1.794	+ .8445		- .4386	+ .1942		1.240
L1,F3	-1.901	+ .9308		+ .00073	+ .3082		2.232
L2,F1	-1.490	+ .5040		- .5775	+ .0659		5.315
L2,F2	- .7925	- .0630	- .1315	+ .1407	+ .0835	+ .2015	6.568
L2,F3	-1.660	+ .8145	- .0783	+ .1132	- .3616	- .2303	7.071

TABLE 1-c Model of Eq.(3); s = 10

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\sigma_w^2 \times 10^{-3}$
L1,F1	- .8923	- .0809		- .0700	- .0817		4.388
L1,F2	-1.801	+ .8517		- .4635	+ .2023		1.213
L1,F3	-1.906	+ .9354		- .0303	+ .3111		2.152
L2,F1	-1.377	+ .3937		- .4764	+ .0698		5.145
L2,F2	- .8812	- .0248	- .0816	+ .0506	+ .0412	+ .2299	6.485
L2,F3	-1.787	+ .9486	- .1168	- .0152	- .4562	- .2703	6.991

TABLE 1-d Model of Eq.(3); s = 15

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\sigma_w^2 \times 10^{-3}$
L1,F1	- .8641	- .1074		- .0428	- .0868		4.365
L1,F2	-1.807	+ .8566		- .4834	+ .1973		1.183
L1,F3	-1.906	+ .9358		- .0318	+ .3100		2.144
L2,F1	-1.502	+ .5152		- .6185	+ .0707		5.088
L2,F2	- .9123	- .0517	+ .0192	+ .0192	+ .0126	_ .2340	6.459
L2,F3	-1.716	+ .8720	- .0934	+ .0597	- .4019	- .2592	6.915

tests performed is not enough, but it certainly establishes a trend to support our argument. Consequently, for s=10 in our case we may conclude that F1 is distinguishable from both F2 and F3 by its $\hat{\phi}_1$, $\hat{\psi}_1$ in L1 and by $\hat{\phi}_2$ and possibly $\hat{\phi}_1$ in L2. We comment again that these conclusions affect different patients differently (different parameters make the same function distinguishable). Therefore, calibration runs must be made on each patient for the functions required and the locations used, all of which depend on the individual patient.

Table 2 AR Parameters; L1,F1

Autoregressive Parameters	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}
s = 3	.8252	.0521	.0904	-	-	-	-	-	-	-	-	-	-	-	-
s = 5	.8293	.0530	.1235	-.0244	-.0154	-	-	-	-	-	-	-	-	-	-
s = 10	.8231	.0562	.1349	-.0286	.0134	-.1151	.0156	.0301	.0396	.0012	-	-	-	-	-
s = 15	.8215	.0559	.2195	-.0291	.0271	-.1216	.0153	.0225	.0463	-.0306	.0601	.0152	-.0485	-.0242	.0302

Table 3 Model for Eq.(3); s = 10; Ranges of Parameters

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\sigma_w^2 \times 10^{-3}$
L1,F1	- .8923 ± .1120	-.0809 ± .1230	-.0700 ± .0880	-.0817 ± .0631	4.388 ± .261
L1,F2	-1.801 ± .0480	+.8517 ± .0160	-.4635 ± .21	-.2023 ± .025	1.213 ± .265
L1,F3	-1.906 ± .010	+.9354 ± .013	-.0303 ± .29	+.3111 ± .016	2.152 ± .69

We may train the amputee to exert always similar force for various functions and control amplitude (at will) independently. Amplitude control is however a subject by itself, and has also other solutions (such as having one separate function for amplitude control which may be toe actuated as in [20]). However, it is interesting to observe that, within a range of amplitudes, parameter patterns for some functions in some locations are sufficiently distinguishable, such that they can be identified. In these cases, σ_w^2 can serve as an amplitude calibrator (see Table 4).

We note that the force/sample-variance relation is not linear and requires calibration, but this does not detract from our argument. In the case of Table 4, the parameter combination pattern at L1 is similar to that of F3, but σ_w^2 makes these distinguishable, noting that for five pounds the pattern of F3 changes significantly (Table 1 gives F3 for one pound). At location L2, function F2 is distinct (at five pounds) from F3 also in the dynamic parameters.

4.2 Models for K-B Filtering (eq. 6)

Following the discussion of Section 3.3, the coefficients for the model of eq. (6), which are equivalent to the K-B filter parameters are given in Table 5.

Note that the orders for F2,F3 at locations L1 and L2 differ for the same reasons as indicated in relation to Table 1 above. Table 5 shows, as expected, that the filter model of eq. (5) may yield easier resolution of functions than does the model of eq. (4). See F2,F3 at L2.

5. Conclusions

In conclusion, it is emphasized that the present study is of initial nature, and the number of signals and patients examined is certainly not sufficient. However, some insight has been given into the stochastic properties of the EMG signals and into the possibilities of time-series analysis of these signals. It is shown that different functions can be resolved from the identified parameters of their ARMA and K-B filter models.

It is noted that the present analysis is only a first step towards a reliable use of the above method in practical EMG control. The problem of resolving several functions being performed simultan-

Table 4 Amplitude Effects (for F2, s = 10)

		$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\psi}_1$	$\hat{\psi}_2$	$\hat{\psi}_3$	$\sigma_w^2 \times 10^{-3}$
L1	1 lb.	-1.801	+.8517	-	-.4635	+.2023	-	1.213
	5 lb.	-1.895	+.9318	-	+.1979	+.3247	-	3.893
L2	1 lb.	- .8812	-.0248	-.0816	+.0506	+.0412	+.2299	6.485
	5 lb.	- .9604	-.5959	+.6223	+.2207	-.1959	+2911	7.835

Table 5 Filtered Model: Parameters of eq.(6); s = 10

	$\hat{\theta}_2$	$\hat{\theta}_3$	$\sigma_u^2 \times 10^{-3}$	$\sigma_v^2 \times 10^{-3}$
L1,F1	+ .4189	-	3.308	.8264
L1,F2	+ .0360	-	1.0806	.1601
L1,F3	+ .0527	-	2.535	.00581
L2,F1	- .2621	-	3.325	.9122
L2,F2	+ .4464	+.07811	5.4705	1.286
L2,F3	2.829	-.5129	42.65	16.18

eously has not yet been attacked. For control purposes the latter problem may be by-passed by performing one function at a time. Avoiding this problem in this way is justifiable in view of the practical constraints of Section 1.1. The problem can be attacked if required by either resolving new parameter combinations, at least for (this must be first shown possible) discrete number of combination levels. The resultant complexity will be proportional to the number of levels thus considered, and which should therefore be kept to a minimum. An approach as in Ref. [19] can also be considered, though it again adds considerable complexity and is time consuming. However, all these considerations are beyond the scope of the present study, and the investigation is still of interest without detracting from the usefulness of the results of this work (especially noting that by-passing the problem is a possibility).

It has been stated that the present analysis considered only stationary EMG signals. An investigation of transients is important. However, its increased complexity justifies the present approach when considering the practical limitations, in prosthesis applications, as stated in Section 1.1. Here we note that the present identification requires only about 150-300 samples to yield adequate results, thus identification can be completed within .15-.25 seconds noting the sampling rate employed (see Sect. 2). Such identification speed may be practical for prosthesis applications with adequate special purpose micro-computer hardware. As long as the model complexity is kept to the one considered in this work, micro-miniaturization via using advanced integrated circuits can facilitate the realization of this identification within cost, weight and volume limits that will allow an amputee to carry the micro-computer in his pocket. The controller itself can thus be of the type discussed in Ref. [20].

We reemphasize that recalibration of the parameter space for different functions must be performed for each patient individually, due to the individual parameter-to-function relations. Also the functions to be used are chosen on the basis of considerations that differ from patient to patient are the electrode locations. Hence, there may or may not be a 1:1 relation between the function actuating the EMG and between the actual prosthesis function (namely if the patient generates a signal relating to a squeeze movement, the control computer may actually actuate a squeeze movement but may alternatively actuate another prosthesis movement, say, an arm bending movement). Obviously, 1:1 relations are preferable from point of view of convenience of the amputee, but make no difference to the computer and controller.

The above discussions point out to considerable further work in identification, recognition and the related control. However, we hope that this work has demonstrated an approach that has statistical rigour (in view of considerable practical constraints) and still has proven to yield results of considerable practical usefulness. Therefore, we hope that we have established a case for further research whose importance is not only beyond statistical analysis and control, but extends also to many other diagnostic EMG applications [23], related to muscular and neural disturbances and disabilities.

References

[1] Mason, C. P.: Practical Problems in Myoelectric Control, Bull. Prosth. Res., No. 13, pp. 39-45, Spring 1970.

[2] Radonjic, D. and Long, C.: Why Myoelectric Control is so Difficult, Advances in External Control of Human Extremities, (Proc. 3rd Internat. Symp. on External Control of Human Extremities) Ed.: M. M. Gavrilovic and A. Bennet Wilson, Belgrade, 1970.

[3] Lyons, C. V.: Multichannel Myoelectric Control, Bull. Prosth. Res., No. 12, pp. 106-117, Fall 1969.

[4] Feeney, R. J. and Hagaeus, I.: Evaluating EMG Advances in External Control of Human Extremities, (Proc. 3rd Internat. Symp. on External Control of Human Extremities) Ed.: M. M. Gavrilovic and A. Bennet Wilson, Belgrade, 1970.

[5] Dorcas, D. S., Dunfield, V. A. and Scott, R. M.: Improved Myoelectric Control System, Med. & Biol. Eng., Vol. 8, pp. 333-341, 1970

[6] Scott, R. N. and Parker, P. A.: Myoelectric Signal Processing, Chapter 7 in: Progress Report 71.1, University of New Brunswick Bio-Eng. Inst., November, 1971.

[7] Kwatny, E., Thomas, D. H. and Kwatny, H. G.: An Application of Signal Processing Techniques to the Study of Myoelectric Signals, IEEE Trans. on Biomed. Eng., Vol. BME-17, pp. 303-312, November, 1970.

[8] Kriefeldt, J. G.: Signal Versus Noise Characteristics of Filtered EMG used as a Control Source, IEEE Trans. on Biomed. Eng., Vol. BME-18, pp. 16-22, January, 1971.

[9] Scott, R. N.: Myoelectric Energy Spectra, Med. & Biol. Eng., Vol. 3, pp. 303-305, 1967.

[10] Zetterberg, L. A.: Estimation of Parameters for a Linear Difference Equation with Application to EEG Analysis, Mathematical Biosciences, Vol. 5, pp. 227-275. 1969.

[11] Fenwick, P., Mitchie, P., Dollmire, J. and Fenton, G.: Mathematical Simulation of the EEG Using an Autoregressive Series, Biomedical Computing, Vol. 2, pp. 281-307, 1971.

[12] Gersch, W.: Status and Prospects of EEG

Spectral Analysis, Proc. 3rd Symp. on Nonlinear Estimation, pp. 287-291, San Diego, 1972.

[13] Graupe, D.: Identification of Systems, Book, Van Nostrand Reinhold Publishing Co., New York, 1972.

[14] Mann, H. B. and Wald, A.: On the Statistical Treatment of Linear Stochastic Difference Equations, Econometrica, 11, pp. 173-220, 1943.

[15] Krause, D. J. and Graupe, D.: Identification of Predictor and Filter Parameter by ARMA (autoregressive-moving-average) Methods, Intl. Jour. Cont., pp. 1021-1027, 1973.

[16] Graupe, D.: Estimation of Upper Bounds of Errors in Identifying AR Models, Proc. 4th Symp. on Nonlinear Estimation Theory, San Diego, 1973.

[17] Saridis, G. N. and Stein, G.: Stochastic Approximation Algorithms for Discrete-Time Systems Identification, IEEE Trans., Vol. AC-13, pp. 515-523, 1968.

[18] Lakacs, E.: Stochastic Convergence, D. C. Heath and Co., Lexington, Massachusetts, 1968.

[19] Graupe, D., Krause, D. J. and Cline, W. K.: Identification of Kalman-Bucy Filters from Noisy Measurement Arrays, Intl. Jour. Sys. Sci., in print.

[20] Graupe, D.: Control of an Artificial Upper Limb in Three Degrees of Freedom, Bull. Prosth. Res., No. 10-17, pp. 25-39, 1972.

[21] Doob, J. L.: Stochastic Processes, John Wiley and Sons, New York, 1953. (Sect. X.1)

[22] Parzen, E.: Stochastic Processes, Holden Day, San Francisco, 1962. (Sect. 3.2)

[23] Marinacci, A. A.: Applied Electromyography, Lea & Febiger, Philadelphia, 1968.

Acknowledgments

Discussions with Dr. E. F. Murphy, Chief, Research and Development Division of the VA Prosthetic and Sensory Aids Service in New York, and of Dr. R. J. Morgan, Electrical Engineering Department, Colorado State University, are gratefully acknowledged.

DEVELOPMENT OF A PROGRAMMABLE, RESPONSIVE CARDIAC PACING SYSTEM

by

Reese S. Terry, Jr.
Cordis Corporation
Miami, Florida 33137

Cordis Omnicor System of programmable cardiac pacing has been designed to meet the following objectives:

(1) Ability to vary output current and pulse rate of an implanted pacer over a wide range, without penetrating the skin.

(2) Simple and reliable programming.

(3) Decreased standby current drain on the pacer battery.

(4) Increased reliability.

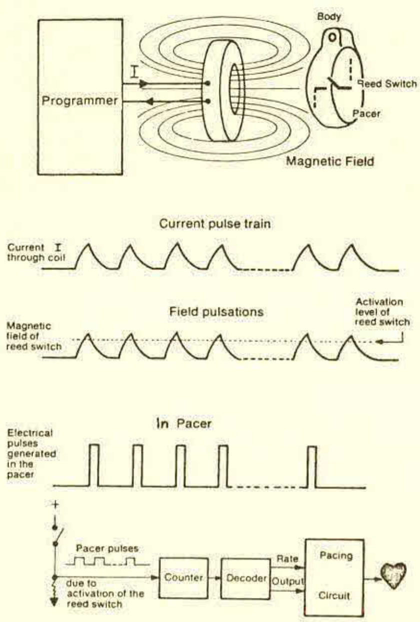
(5) Decreased volume of electronics package.

(6) Retention of reasonable cost.

The Omnicor System consists of an implantable pacer with integrated circuits and an external, battery powered programmer.

The pacer can be programmed at any of six rates and at any of four output current levels.

To program the pacer, the physician places the programmer over the implanted pacer, dials the desired rate and current, and presses the program button.



A 330 Hz magnetic pulse train was chosen as the communication link to a magnetic reed switch in the pacer.

The probability of accidental programming is extremely low. Environmental magnetic fields of the strength and frequency required to program have not been discovered.

The hand-held programmer transmits a pulse train of high-intensity magnetic pulses. Each magnetic pulse changes the pacing rate, and every eight pulses change the output current level. Thus, the total is a multiple of the two functions.

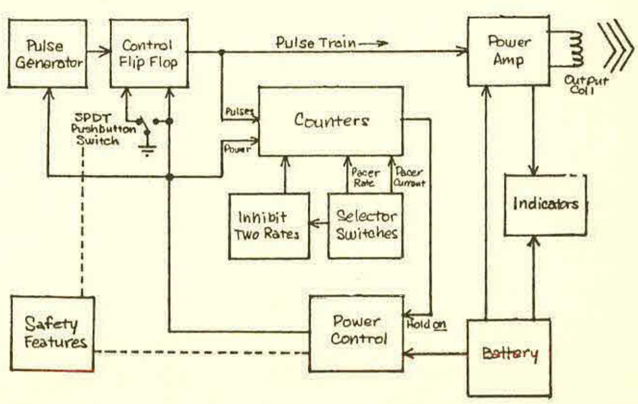
The coding was designed in this manner so that a possible error in programming could be easily detected as an incorrect rate on an electrocardiograph machine.

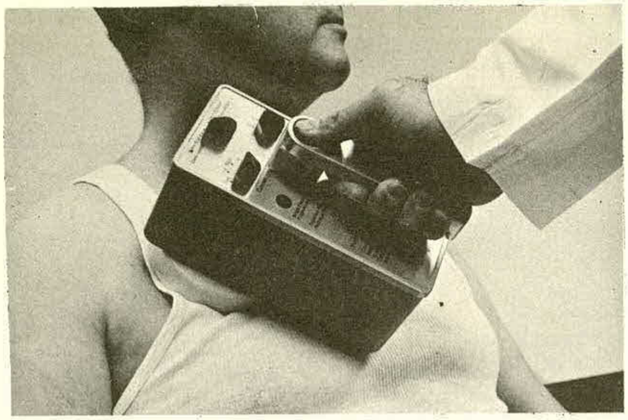
The pacer decoder input is a wave-shaping filter network which discriminates against slower magnetic signals, such as 120 Hz and direct current magnetic fields.

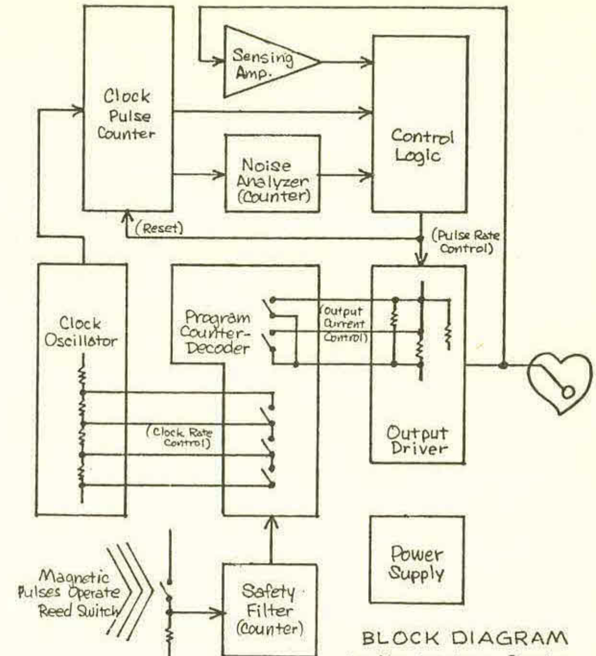
The decoder must receive eight pulses at 330 Hz to enable programming. Each additional pulse is counted by a binary counter which controls solid state switching.

This switching, within the decoder section of the pacer, shorts or connects resistances in ladder networks, thereby controlling the clock rate, the pacing rate and the output current level of the implanted pacer.

BLOCK DIAGRAM
Omnicor Programmer
Model 166-B







BLOCK DIAGRAM
Cordis Omnicor Series
Programmable Pacer

Pacing circuitry consists of a COS/MOS monolithic, integrated circuit which includes a clock multivibrator running at 512 times the output pacing rate, a nine-stage counter and logic gates to set the pulse width and refractory time.

The integrated circuit also incorporates a monolithic, linear sensing amplifier to detect electrical signals of the heart; and a noise analyzer.

Four modes of pacing are achieved with the one logic chip: ventricular synchronous, ventricular inhibited, atrial synchronous and asynchronous. The mode of the pacer is fixed during manufacture.

The interval between clock oscillator pulses is utilized to fix the width of the output pulse.

The refractory (insensitive) segment for all heart-responsive modes is three-eighths of the pacing cycle, or 192 counts.

In the ventricular-synchronous mode, an R-wave inhibited segment follows immediately after the refractory segment. The purpose of the inhibited segment, which is one-eighth of the pacing cycle, is to provide a sufficiently long alert segment (five-eighths of the pacing cycle) but to prevent noise from causing the pacer to stimulate during the vulnerable part of the heart cycle.

In the three modes which are responsive to the heart, Omnicor pacers use a unique method of testing for electrical noise, and respond by reverting to asynchronous operation when noise makes accurate sensing of heartbeats impossible.

A counter is enabled during the final one-sixth of the refractory segment. If four or more noise pulses are received from the amplifier during the testing segment, the pacer remains refractory until the end of the pacing cycle and a pacing pulse is issued.

COS/MOS logic circuits in the decoder (the programmable portion of the pacer) account for less than a tenth of one percent of the total battery drain. The total standby current drain of the pacer has been reduced to six microamperes by use of integrated circuits in place of discrete components.

Thick-film, hybrid circuit techniques are used to assemble the integrated circuit and capacitor chips. All active circuitry is hermetically sealed in two metal packages. External thick-film resistors, which are insensitive to moisture, are trimmed to the operating parameters specified for the device. The volume required for the electronics has been reduced to one cubic inch.

Since devices of the COS/MOS logic family can operate in a wide range of voltages, the pacer can be programmed after the loss of power supply voltage equivalent to that of two of the five cells in the pacer battery pack.

Impending battery depletion is indicated by a decrease in fixed rate. Retention of capture during voltage decline is maintained because the duration of the output pulse increases in proportion to the decrease in intensity of the output pulse.

By programming the output current immediately above the lowest setting that maintains capture, the lifetime of the implanted pacer can safely be maximized for given patient conditions and battery capability.

COMPUTER IMPLEMENTATION

OF A

PARAMETRIC MODEL FOR BIPED LOCOMOTION KINEMATICS

By

Thomas Charles Hartrum, Ph.D.*

Abstract

Although many different aspects of locomotion have been studied, no one has combined all gait aspects into a complete mathematical model of locomotion. The primary contribution of this paper is the parametric representation of the kinematic motion of biped locomotion incorporating all aspects of biped gait, and a computer implementation of this representation. First the human body is represented by a series of connected links. A set of body variables, for example joint angles or joint coordinates, is chosen to form a set of basis variables. Most of these variables are dependent only on time during the entire gait cycle and are called independent variables. During that portion of the gait cycle when both feet are on the ground, more constraints are added to the body, and some of these otherwise independent variables become dependent on other variables. Thus these so-called semi-independent variables are independent during only a part of the gait cycle. Experimental data for these independent and semi-independent variables has been recorded from actual human beings by other investigators. Their resulting curves are approximated here by first order sinusoidal functions. The magnitudes and phases of these functions, along with the physical dimensions of the body, form a gait vector whose components are 101 gait parameters. These 101 parameters completely specify the gait for this model. In particular, they specify the independent and semi-independent variables. Next the three-dimensional coordinates of all the body joints are calculated as a function of these variables. A computer implementation of this model is then used which allows either printed data of body joint coordinates and angles, or a pictorial representation of a figure walking using the computer graphic display

---

*This work was supported in part by the National Science Foundation under Grant GK-25292, at The Ohio State University, Columbus, Ohio, 43210. The author is currently assigned to the U.S. Army Electronics Command, with home address 261 B Eaton Crest Drive, Eatontown, N. J. 07724.

unit. The resulting motion for normal locomotion is presented and compared with experimental data. Finally, some consideration is given to the simulation of pathological gait with this model.

1. INTRODUCTION

The art of walking, taken for granted by most people, has never been duplicated mechanically. The potential applications of artificially simulating the walking process are many and varied. Robots could replace men for carrying heavy loads or for working in hostile environments. The simulation of locomotion could provide a basis for the design of better protheses and orthoses. A better understanding of gait disorders might result, which could lead to computerized analysis of pathological gaits. This could provide a tool for use by the physician for diagnosing gait disorders. Although many different aspects of locomotion have been studied, no one has before combined all gait aspects into a complete mathematical model of locomotion. This author [1] has developed a parametric representation of the kinematic motion of biped locomotion incorporating all aspects of biped gait. A computer program is used to implement the equations of this model.

In developing a locomotion model, two fundamental questions need to be answered. First, what features need to be included ? Secondly, how are these features to be simulated ? The answers to these questions are based on the work of two authors. First, J. B. Saunders, et al, [2] have defined six "determinants," or features, of gait. These are (1) pelvic rotation, defined as the rotation of the pelvis in a horizontal plane about a vertical axis; (2) pelvic tilt, the rotation of the pelvis in the frontal plane about a horizontal axis; (3) knee flexion in the stance phase, which is the bending of the knee of the weight-bearing leg as the body passes over that foot; (4 & 5) foot and knee mechanisms, the combined action of the foot rotating first around the ankle joint, then around the ball of the foot, and the flexing and locking of the knee of the weight-bearing leg; (6) lateral displacement of the pelvis, the side-to-side motion of the hips during forward locomotion. In simulating these six determinants of gait, the data of M. P. Murray, et al, [3,4] is used. This data is presented as time-dependent curves of the motion of many of the body angles and the coordinates of the joints. These curves show the need for other features in the model beyond the six gait determinants of J. B. Saunders, et al, [2] . These include leg motion during the swing phase; foot angle, measured between the foot on the ground and the direction of travel; base width, the lateral distance between the feet; and upper extremity motion, including both forward and lateral tipping of the trunk, chest rotation, and arm movements. M. P. Murray [3,4] presents data for more variables than are needed to completely specify body motion. Thus one set of these variables can be simulated as time-dependent functions, and the remainder of them used to evaluate the resulting motion of the model.

In this paper the physical model is first determined. Then a choice of variables is made which will specify the motion of the model. These variables are then simulated as a function of time based on the data of M. P. Murray [3,4] . The coordinates of all the body joints are then calculated as a function of these time-dependent variables. Finally, a computer simulation is used to implement these equations, with pictorial output on a graphic display terminal of the model walking. The magnitudes and phases of the time-dependent variables, along with the physical dimensions of the body, form a gait vector of dimension 101. These 101 gait parameters completely specify the motion of the model. A discussion of the parameter values for normal locomotion, and a comparison of the resulting motion with experimental data are presented. Some consideration is also given to the pathological capabilities of the model.

2. DERIVATION OF A KINEMATIC MODEL

2.1 PARAMETRIC SIMULATION OF INDEPENDENT VARIABLES

The physical model is a stick figure of sixteen straight-line links. These are the toes, feet, lower legs, upper legs, pelvis, trunk, neck, shoulder link, upper arms, and lower arms. The toe, foot, lower leg, and upper leg are constrained to be coplaner. The heels are separated laterally by a fixed distance B, known as the base. The hip joints are separated laterally by a distance W. The resulting pelvis is a straight line of varying length between these hip joints. Specifying the lateral separation rather than the actual length of the pelvic link is equivalent to a small angle approximation for the pelvic angles, and simplifies the equations somewhat. The trunk is connected to the center of the pelvic link. At the top of the trunk is a short neck link which is always vertical. The shoulder link is also connected at this point. It can

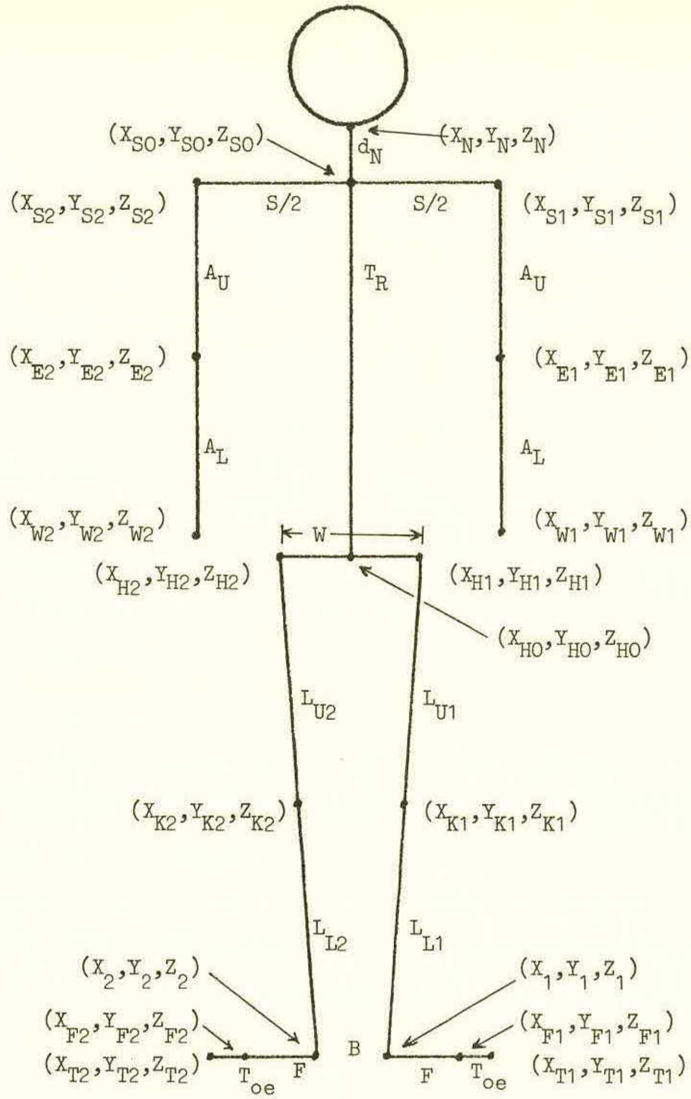


Fig. 1. The physical model of MAN

rotate about this point, but it is constrained to a horizontal plane. The arms are connected to the ends of the shoulders and are constrained to move in vertical planes. Figure 1 shows the resulting physical model.

At any given instant of time the position of the body can be completely described by specifying a set of variables. One such set would be the coordinates of all the joints. Both for comparison with available data and for consideration of future instrumentation, however, the angles between body segments form a more practical set in many cases. The chosen set forms an independent basis for describing the body, and these variables are termed "independent variables." The remaining variables (coordinates and angles) are determined by these independent variables, and are called "dependent variables." During different portions of the gait cycle the number of constraints on the system varies. For example, there are less constraints when one foot is raised than when both feet are on the ground. Thus it is necessary for some variables to be dependent during part of the cycle and independent the rest of the time. These variables are called "semi-independent variables." The primary basis for choosing independent variables is the availability of existing data, such as presented by M. P. Murray [3,4] . The second basis for selection is a consideration of possible instrumentation. Thus angles tend to be chosen instead of linear coordinates. The selected independent variables are the hip angle θ_{Hi}, the knee angle θ_{Ki}, and lateral displacement of the pelvis. Semi-independent variables are ankle angle θ_{Ai}, pelvic tilt θ_P, and pelvic rotation θ_R.

In order for the figure to move, the independent variables (including the independent portion of the semi-independent variables) must be expressed as mathematical functions of time. In order to best simulate real biped locomotion, these time functions are formed by mathematical approximations of actual data measured from walking humans, primarily the data of M. P. Murray [3,4] . As an example, consider the general waveform for the hip angle θ_{H1} as shown in Figure 2. Since this is a periodic function with period P, it could be represented as an infinite series with period P. However, if the gait cycle is divided into subintervals at the switching times $\tau_{D,1}$, $\tau_{5,1}$, and $\tau_{9,1}$, the function can be represented separately within each of these subintervals instead of over the whole cycle. Closer examination of Figure 2 indicates that within each subinterval the function can be represented by a simple first-order sinusoid instead of an infinite series. When all of the data of M. P. Murray [3,4] is considered in this light, it appears that the various curves can all be represented by first and second order functions in subintervals of the gait cycle.

Since these subintervals are chosen to coincide with extrema of the sinsusoidal curves, there is no phase information necessary. All that is needed to describe the function is the period of each sinusoid and its magnitude. The period can be determined by specifying the subinterval boundaries as a percentage of the total gait cycle. These "switching times" are the values such as $\tau_{5,1}$ shown in Figure 2. The magnitude can be determined by specifying the maximum and minimum values which occur at the end points of the subinterval. These "magnitudes" for the left hip would be the values of θ_{H1min}, θ_{H1dip}, θ_{H1max}, and θ_{H1over} as shown in Figure 2. What is of importance here is that not only can these curves be represented by a finite number of functions instead of an infinite series, but also they can be specified by a set of magnitudes and switching times that have physical significance instead of a set of meaningless coefficients of a series expansion. These magnitudes and switching times, together with the physical dimensions of the body, are defined as "gait parameters." The total set of all gait parameters needed to specify the motion of the body during locomotion is called the "gait vector γ ."

The hip angle shown in Figure 2 is made up of four first-order sinusoids. For example, for $\tau_{D,1}P < t < \tau_{5,1}P$

$$\theta_{H1} = \tfrac{1}{2}(\theta_{H1max} + \theta_{H1dip}) - \tfrac{1}{2}(\theta_{H1max} - \theta_{H1dip}) \cdot \cos\left[\frac{\pi(t - \tau_{D,1}P)}{(\tau_{5,1} - \tau_{D,1})P}\right] \quad (1)$$

This is for the left leg. For all of the variables the right leg

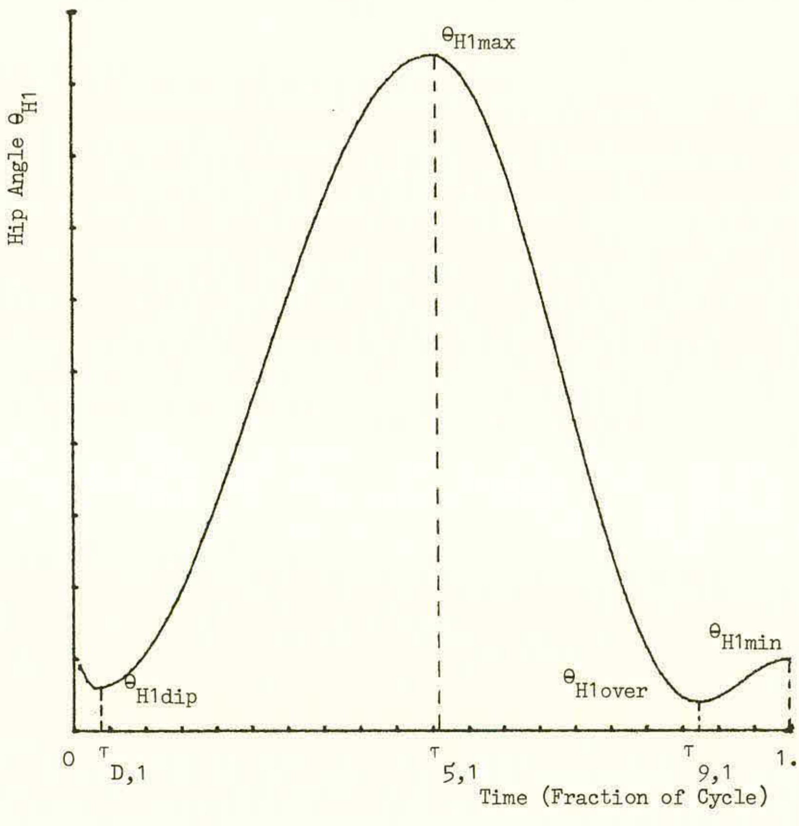


Fig. 2. Waveform for the hip angle of the left leg.

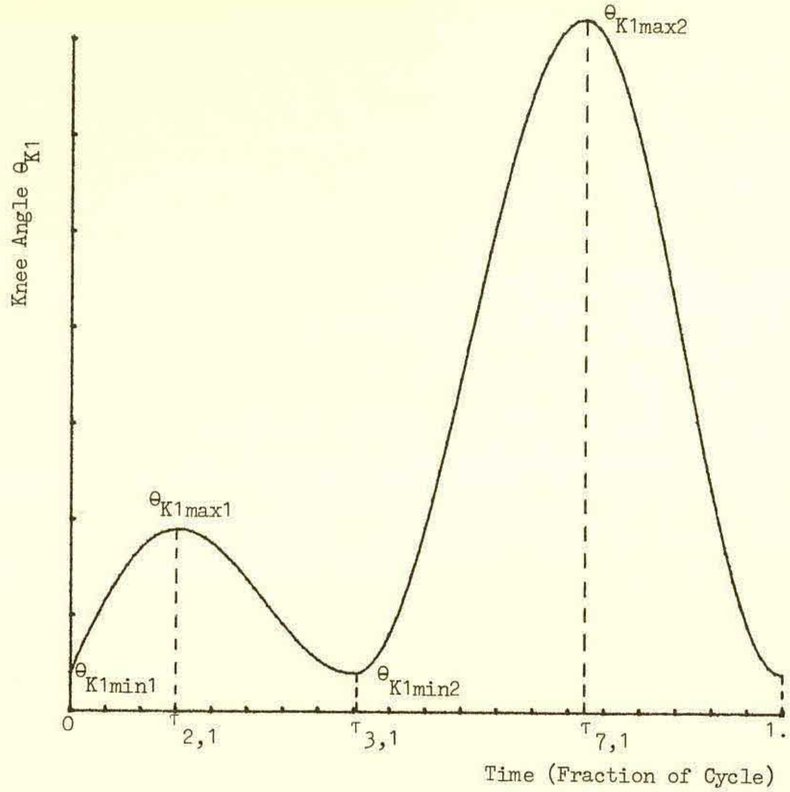


Fig. 3. Waveform for the knee angle of the left leg.

equations are similar, but are delayed by an amount ΦP, where Φ is the phase of the right foot with respect to the left foot. The value of Φ is a gait parameter. The contribution to the gait vector due to the hip angles is the set (θ_{Himin}, θ_{Hidip}, θ_{Himax}, θ_{Hiover}, $\tau_{D,i}$, $\tau_{5,i}$, $\tau_{9,i}$) for $i = 1,2$. The knee angle θ_{K1} is shown in Figure 3. It is also made up of four sinusoids. The contribution to the gait vector due to the knee angles is the set (θ_{Kimin1}, θ_{Kimax1}, θ_{Kimin2}, θ_{Kimax2}, $\tau_{2,i}$, $\tau_{3,i}$, $\tau_{7,i}$) for $i = 1,2$. In Figure 4 the waveform for lateral pelvic displacement Y_{HO} is shown. This is made up of two simple sinusoids, and its contribution to the gait vector is the set (Y_{HOmin}, Y_{HOmax}, $\tau_{H,1}$, $\tau_{H,2}$). These are the three primary independent variables.

The first semi-independent variable to be considered is the ankle angle θ_{A1} as shown in Figure 5. During that portion of the cycle between $\tau_{1,1}$ and $\tau_{4,1}$ the ankle angle is completely determined by the hip and knee angle functions. Thus the values for $\theta_{A1T1,1}$ and $\theta_{A1T4,1}$ are determined and serve as boundary values for the independent functions in the adjacent subintervals. The first and last segments of the curve of Figure 5 are linear functions, while the remaining two independent segments are quadratic functions. The total contribution to the gait vector from the ankle angles is the set ($\theta_{AiT0,i}$, $\theta_{AiT6,i}$, $\theta_{AiT8,i}$, $\tau_{1,i}$, $\tau_{4,i}$, $\tau_{6,i}$, $\tau_{8,i}$) for $i = 1,2$. The next semi-independent variable, shown in Figure 6, is pelvic tilt θ_P. During those portions of the gait cycle when both feet are in contact with the ground (known as the double stance phase), the locations of the two hip joints are completely determined. Therefore the angle of pelvic tilt is determined, and θ_P is a dependent variable as shown by the dashed portions of Figure 6. The values $\theta_{Pinitial1}$, $\theta_{Pfinal1}$, $\theta_{Pinitial2}$, $\theta_{Pfinal2}$ are specified boundary conditions. During the swing phase for each leg (when the corresponding foot is off the ground), pelvic tilt θ_P is an independent variable. Consider the swing phase of leg 1, for $\tau_{6,1}P < t < P$. Since the extremum can occur at any time during this interval, both magnitude and phase information is needed. Therefore θ_P can be written

$$\theta_P = \theta_{Pav} + A \cdot \sin\left(\frac{2\pi t}{P} + \Psi_P\right) \qquad (2)$$

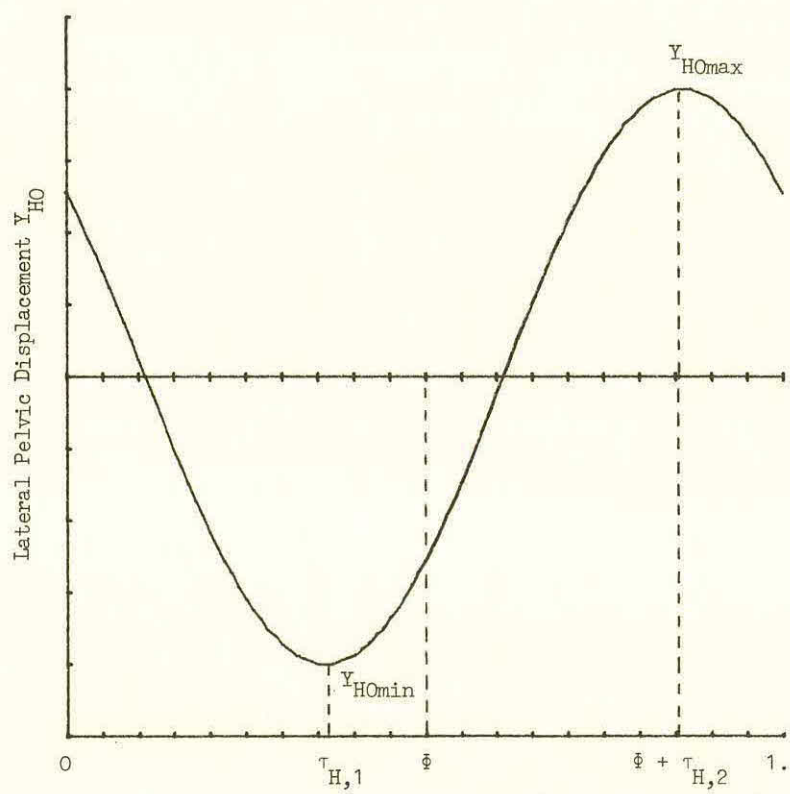


Fig. 4. Waveform for lateral displacement of the pelvis.

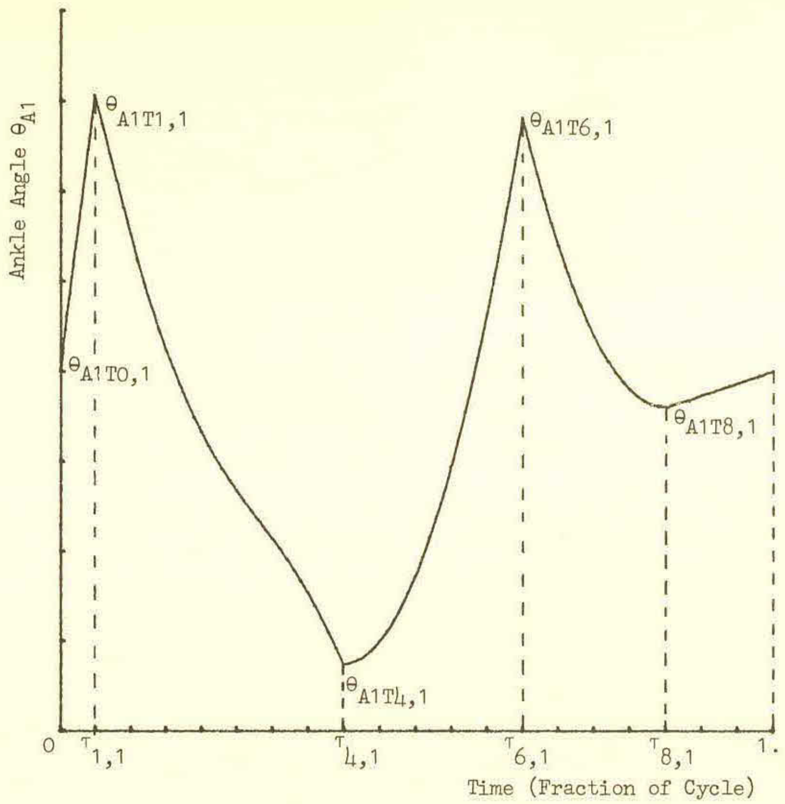


Fig. 5. Waveform for the ankle angle of the left leg.

where θ_{Pav}, A, and Ψ_P are unknowns. The three known conditions required to solve this are the initial and final values, and the minimum value θ_{Pmin} which is a gait parameter. Since the value of the sine at θ_{Pmin} is -1, then at that point equation (2) becomes

$$\theta_{Pmin} = \theta_{Pav} - A \tag{3}$$

Solving for A and substituting into equation (2) yields

$$\theta_P = \theta_{Pav} + (\theta_{Pav} - \theta_{Pmin})\cdot\sin\left(\frac{2\pi t}{P} + \Psi_P\right) \tag{4}$$

Next the boundary values are used to obtain

$$\theta_{Pfinal1} = \theta_{Pav} + (\theta_{Pav} - \theta_{Pmin})\cdot\sin(\Psi_P) \tag{5}$$

and

$$\theta_{Pinitial1} = \theta_{Pav} + (\theta_{Pav} - \theta_{Pmin})\cdot\sin(2\pi\,\tau_{6,1} + \Psi_P) \tag{6}$$

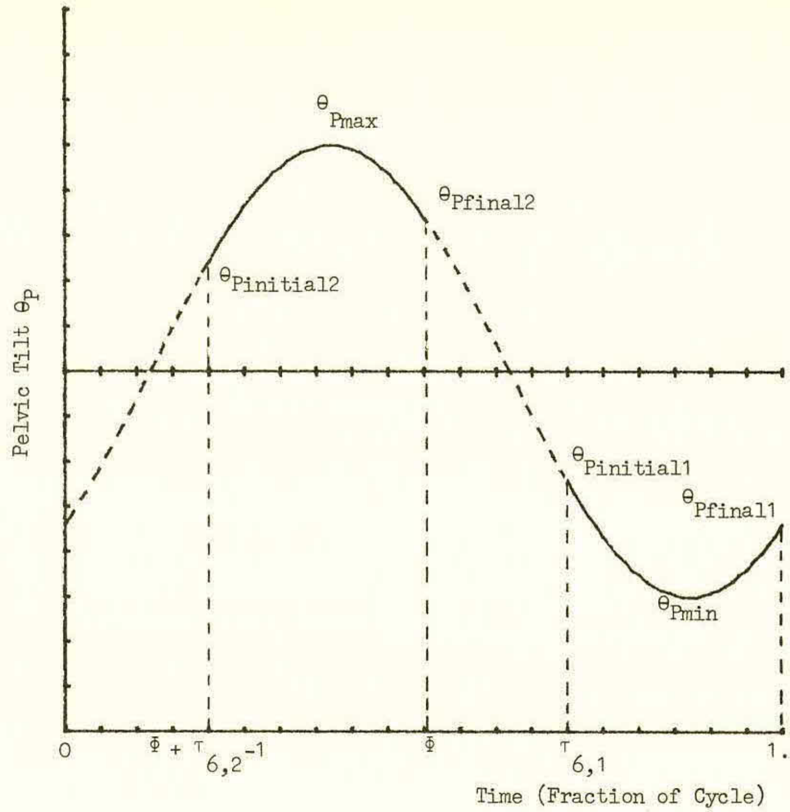


Fig. 6. Waveform for pelvic tilt angle θ_P.

Equations (5) and (6) can be solved for θ_{Pav} and Ψ_P. The resulting equations are

$$\theta_{Pav} = \frac{\theta_{Pfinal1} + \theta_{Pmin}\cdot\sin(\Psi_P)}{1+\sin(\Psi_P)} \tag{7}$$

and

$$\left[-\theta_{Pinitial1} + \theta_{Pfinal1} + (\theta_{Pfinal1} - \theta_{Pmin})\cdot\cos(2\pi\tau_{6,1})\right]\ \sin(\Psi_P)$$

$$+\left[(\theta_{Pfinal1} - \theta_{Pmin})\cdot\sin(2\pi\tau_{6,1})\right]\ \cos(\Psi_P)$$

$$= \theta_{Pinitial1} - \theta_{Pfinal1} \tag{8}$$

Equation (8) is of the form

$$\alpha\sin(x) + \beta\cos(x) = \gamma \tag{9}$$

whose solution is

$$x = \sin^{-1}\left(\frac{\gamma}{\sqrt{\alpha^2+\beta^2}}\right) - \tan^{-1}\left(\frac{\beta}{\alpha}\right) \tag{10}$$

Due to the different quadrants in which $\tan^{-1}(\beta/\alpha)$ can lie, and due to restrictions on the range of the inverse functions on some computers, it is necessary to replace γ with $-\gamma$ if α is negative. Therefore

$$\gamma = \theta_{Pinitial1} - \theta_{Pfinal1}$$

$$\text{for} \quad -\theta_{Pinitial1} + \theta_{Pmin} + (\theta_{Pfinal1} - \theta_{Pmin})\cdot\cos(2\pi\tau_{6,1}) > 0 \tag{11}$$

and

$$\gamma = \theta_{Pfinal1} - \theta_{Pinitial1}$$

$$\text{for} \quad -\theta_{Pinitial1} + \theta_{Pmin} + (\theta_{Pfinal1} - \theta_{Pmin})\cdot\cos(2\pi\tau_{6,1}) < 0 \tag{12}$$

Then

$$\Psi_P = \sin^{-1}\left\{ \gamma \Big/ \sqrt{\left[-\theta_{Pinitial1} + \theta_{Pmin} + (\theta_{Pfinal1} - \theta_{Pmin})\cos(2\pi\tau_{6,1})\right]^2 + \left[(\theta_{Pfinal1} - \theta_{Pmin})\sin(2\pi\tau_{6,1})\right]^2} \right\} - \tan^{-1}\left\{ \frac{(\theta_{Pfinal1} - \theta_{Pmin})\sin(2\pi\tau_{6,1})}{\left[-\theta_{Pinitial1} + \theta_{Pmin} + (\theta_{Pfinal1} - \theta_{Pmin})\cdot\cos(2\pi\tau_{6,1})\right]} \right\} \tag{13}$$

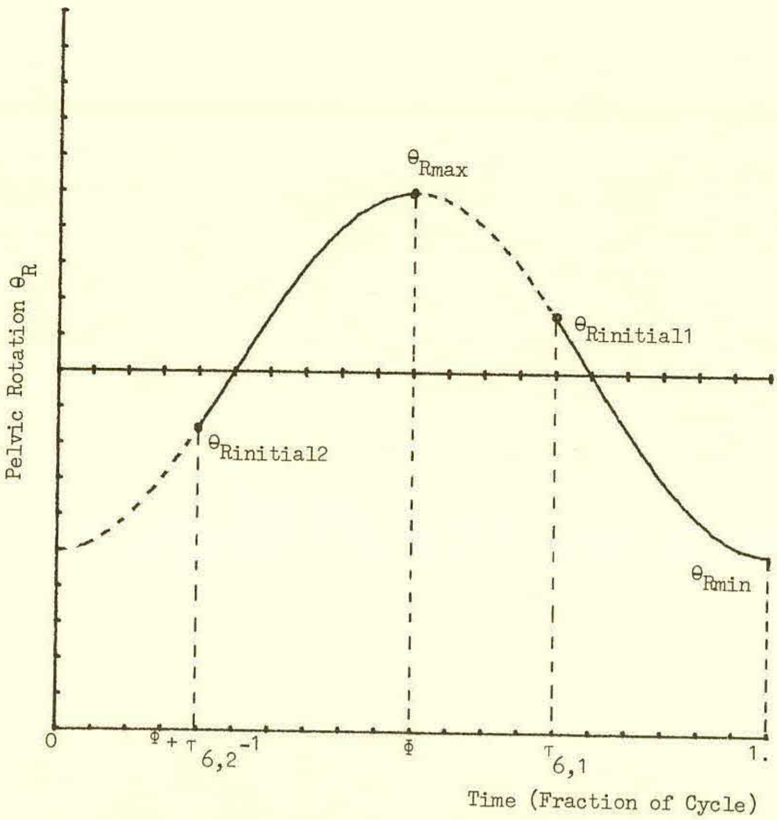


Fig. 7. Waveform for pelvic rotation angle θ_R.

This value for Ψ_P can then be substituted into equation (7) to calculate θ_{Pav}. This specifies θ_P during the swing phase of leg 1. A similar set of equations is used for the swing phase of leg 2. The contribution to the gait vector from pelvic tilt is the set of magnitudes $(\theta_{Pmin}, \theta_{Pmax})$.

The remaining semi-independent variable is pelvic rotation θ_R, shown in Figure 7. The situation for θ_R is similar to θ_P. However, the extreme values occur at the end point of each swing phase, so that no phase needs to be calculated. For the swing phase of leg 1 the angle θ_R can be written

$$\theta_R = \theta_{Rmin} + (\theta_{Rinitial1} - \theta_{Rmin}) \left[\frac{1-\cos(2\pi t / P)}{1-\cos(2\pi \tau_{6,1})} \right] \tag{14}$$

A similar equation can be written for the swing phase of leg 2. The contribution to the gait vector due to pelvic rotation is the set of magnitudes $(\theta_{Rmin}, \theta_{Rmax})$. All of the independent and semi-independent variables have now been completely specified by a set of gait parameters.

2.2 KINEMATIC EQUATIONS OF MOTION

To describe the instantaneous position of the body the standard aircraft coordinate system is used. The X-axis points in the direction of travel, the Z-axis points vertically downward, and the Y-axis points to the right. The origin of this system is located on the ground such that the X-axis lies midway between the Y coordinates of the two heels. Thus the origin is centered below the figure when it is standing at attention. When the figure is moving, the coordinate system moves forward with constant velocity v equal to the average velocity of the figure. The upper leg, lower leg, foot, and toe are coplaner, as shown in Figure 8. The hip, knee, and ankle angles are measured in the plane of the leg. If a new coordinate system (X''', Y''', Z''') is defined with its origin at the ball of the foot and the $X''' - Z'''$ plane coinciding with the plane of the leg, as shown in Figure 8, then the two coordinate systems can be related as follows.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = T^{-1} \begin{bmatrix} X''' \\ Y''' \\ Z''' \end{bmatrix} + \begin{bmatrix} X_{F1} \\ Y_{F1} \\ Z_{F1} \end{bmatrix} \tag{15}$$

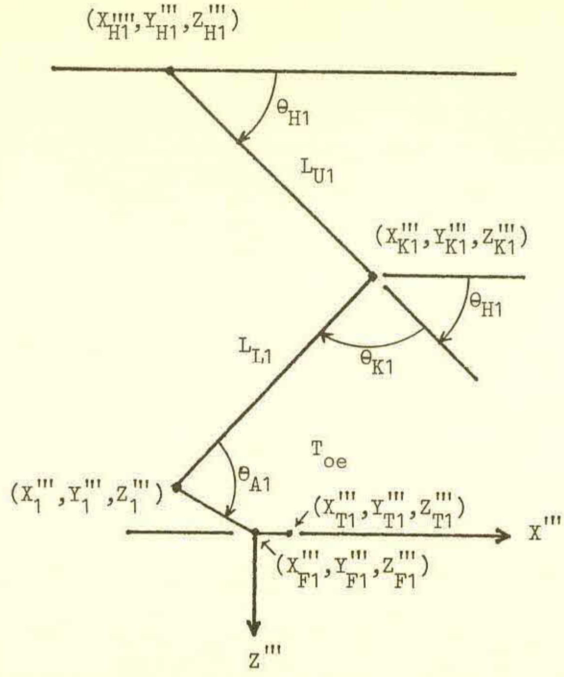


Fig. 8.-- Leg in the (X''', Y''', Z''') system.

where

$$T^{-1} = \begin{bmatrix} \cos(\theta_{F1}) & \sin(\theta_{F1})\cos(\theta_{L1}) & -\sin(\theta_{F1})\sin(\theta_{L1}) \\ -\sin(\theta_{F1}) & \cos(\theta_{F1})\cos(\theta_{L1}) & -\cos(\theta_{F1})\sin(\theta_{L1}) \\ 0 & \sin(\theta_{L1}) & \cos(\theta_{L1}) \end{bmatrix} \quad (16)$$

In equation (16) the angle θ_{Fi} $(i = 1,2)$ is the angle between the foot when it is flat on the ground and the X-axis. The values of θ_{F1} and θ_{F2} are gait parameters. The angle θ_{Li} $(i = 1,2)$ is the angle between the plane of the leg and the horizontal X-Y plane. The derivation of an expression for θ_{Li} will be discussed below. In equation (15) the coordinates (X_{F1}, Y_{F1}, Z_{F1}) are the position of the ball of the foot in terms of the (X, Y, Z) system. Here Z_{F1} is zero since the foot is on the ground, $Y_{F1} = -\frac{1}{2}B - F\cdot\sin(\theta_{F1})$ where B is the lateral distance between the heels and F is the length of the foot, and $X_{F1} = \gamma_1 + F.\cos(\theta_{F1}) - vt$ where γ_1 is the initial position of the foot (derived later) and v is the average velocity of the body as described above. As can be seen from Figure 8, the coordinates of the leg joints can be written easily in the (X''', Y''', Z''') system. Then they can be transformed into the (X, Y, Z) system using equation (15). This is done for all the joints in Figure 8. One expression is of particular importance, that for Y_{H1}.

After equation (15) is applied, Y_{H1} is a function of the unknown θ_{L1}. But since the lateral distance between the hips is a constant W, then $Y_{H1} = Y_{HO} - \frac{1}{2}W$, where Y_{HO} is the midpoint of the pelvis. But Y_{HO} is an independent variable. Therefore Y_{H1} is known as a function of time. The coordinate equation for Y_{H1} can be solved to give an expression for θ_{L1}. This discussion, as can be seen from Figure 8, is valid only when the ball of the foot is on the ground. Similar equations can be derived for those times when only the heel is on the ground, and when the foot is off the ground. Then the motion of the leg is described for all of the gait cycle. During the stance phase it is a function of the constants Y_i (the initial foot position) and v (the average velocity of the figure). The initial position of foot 1, Y_1, is calculated by assuming that the coordinate origin is located directly below the center of the pelvis at $t = 0$ and calculating the position of the left foot. Next the stride length of each foot is calculated. These are added together and divided by the length P of the gait cycle to determine the average velocity v. Then the initial position of foot 2 can be found by calculating the location of the right foot at $t = \Phi P$. This completely specifies the motion of each leg during stance phase.

During swing phase the leg is specified with respect to the hip joint position. But the hip joint of the swinging leg is easily related to the hip joint of the stance leg by the angles of pelvic tilt and pelvic rotation, and by the pelvic width W. Since θ_P and θ_R are independent variables during the swing phase, the lower extremity motion is thus specified for the entire gait cycle P.

The upper extremities are controlled by a set of independent variable angles. The magnitudes and switching times of these angles are all elements of the gait vector. The motion of the trunk has two angles associated with it. First is trunk leaning θ_L. This simulates the side-to-side leaning of the trunk, with one excursion to each side during one gait cycle. The trunk tipping angle θ_T represents the forward and backward motion of the trunk. The trunk tips forward twice during each gait cycle (once for each heel strike) and rearward twice. The shoulder link rotates in a horizontal plane by an angle θ_C, the chest angle. The upper arms are controlled at the shoulder joints by the shoulder angles θ_{Si} $(i = 1,2)$, and the lower arms are moved by the elbow angles θ_{Ei} $(i = 1,2)$. It should be noted that although upper extremity motion may be important for balance in actual locomotion, for the kinematic model the upper extremity motion is completely independent of lower extremity motion.

3. IMPLEMENTATION AND RESULTS

The computer simulation MAN uses the equations described in the preceding section to simulate the motion of a biped during locomotion, and to output the results in several useful forms. Any of the 101 gait parameters can be varied. The output can be printed data for many of the body angles and coordinates. Alternately, a computer graphic display terminal can be used to display a walking stick figure which can be viewed from the side, front, top, or obliquely. Or the graphic display can be used to view the time-dependent graph of many of the body variables.

The parameter values for the physical dimensions of the body are based on a diagram by R. Drillis and R. Contini [5]. The remaining parameter values are based on data by M. P. Murray [3,4], R. Beckett and K. Chang [6], and D. H. Jacobson and C. K. Chow [7]. Where data conflicted, that of M. P. Murray [3,4] usually prevailed. The results for the independent variables matched the data of M. P. Murray [3,4] very well, as seen for hip, knee, and ankle angles in Figures 9, 10, and 11 respectively. However, some of the dependent variables were less than satisfactory. The 101 gait parameters were then manually tuned one at a time to achieve a satisfactory set of results. Graphs for pelvic tilt, pelvic rotation, and the vertical path of the heel are shown in Figures 12, 13, and 14 respectively.

As can be seen from Figures 9 through 14, a reasonably close representation of normal locomotion can be attained with the model. The model is designed to allow for unsymmetrical motion so that it can be used for some pathological gaits. A complete analysis of this capability is impossible due to the lack of data from pathological cases. However, M. P. Murray [3] does present some limited data which permits some consideration of the pathological capability of the model. Figures 15, 16, and 17 show the simulation of the hip, knee, and ankle angles for a patient with bilateral residuals of advanced paralysis agitans.

4. SUMMARY

The MAN simulation provides a fairly complete model of human locomotion in that all of the primary features of human locomotion are represented. The physical model consists of straight rigid links for most of the body segments. These links can be changed in length, and the leg and foot lengths can be independently changed for each side, allowing for unsymmetrical body structure. The hip angles, knee angles, and lateral body motion are independent variables, whose magnitudes and switching times are controllable gait parameters. Ankle angle, pelvic tilt, and pelvic rotation are independent during parts of the gait cycle, and their magnitudes and switching times also contribute to the gait parameters. These variables all combine to provide simulation of the six determinants of gait of J. B. Saunders, et al [2]. Combined with the upper extremities, 101 gait parameters completely specify the motion of the human body during the locomotion cycle. Some capability for pathological gait simulation is also possible with this model.

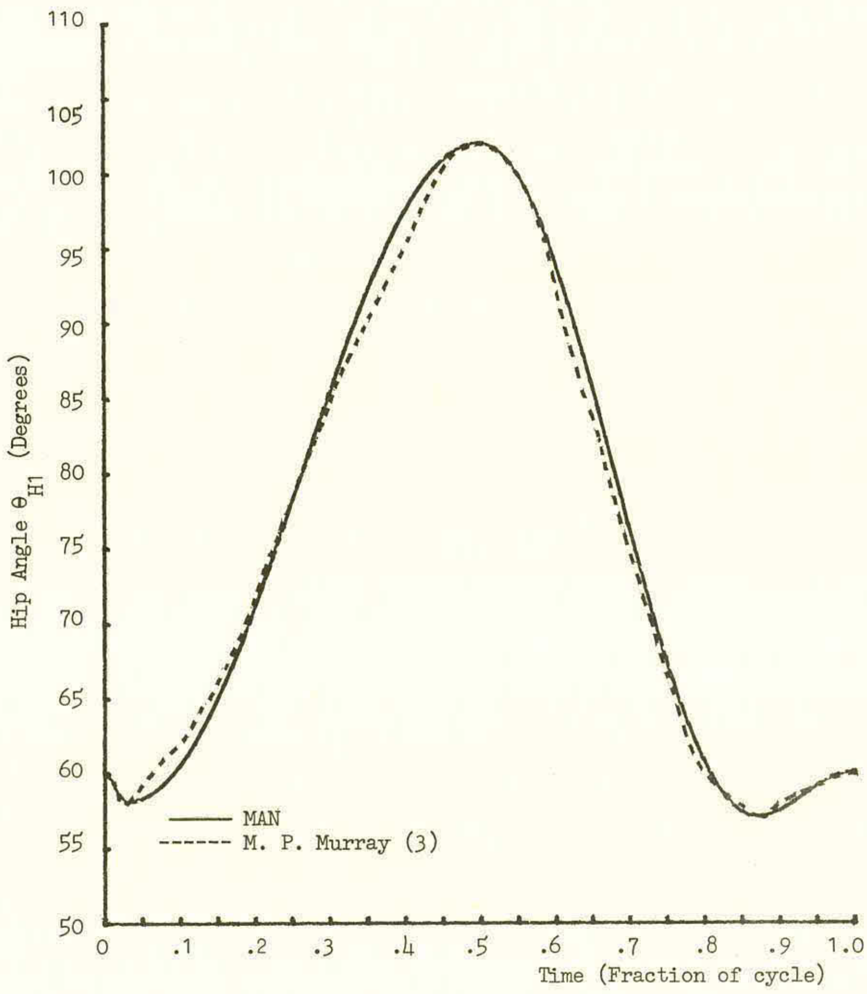


Fig. 9. Hip angle results for initial set of gait parameters.

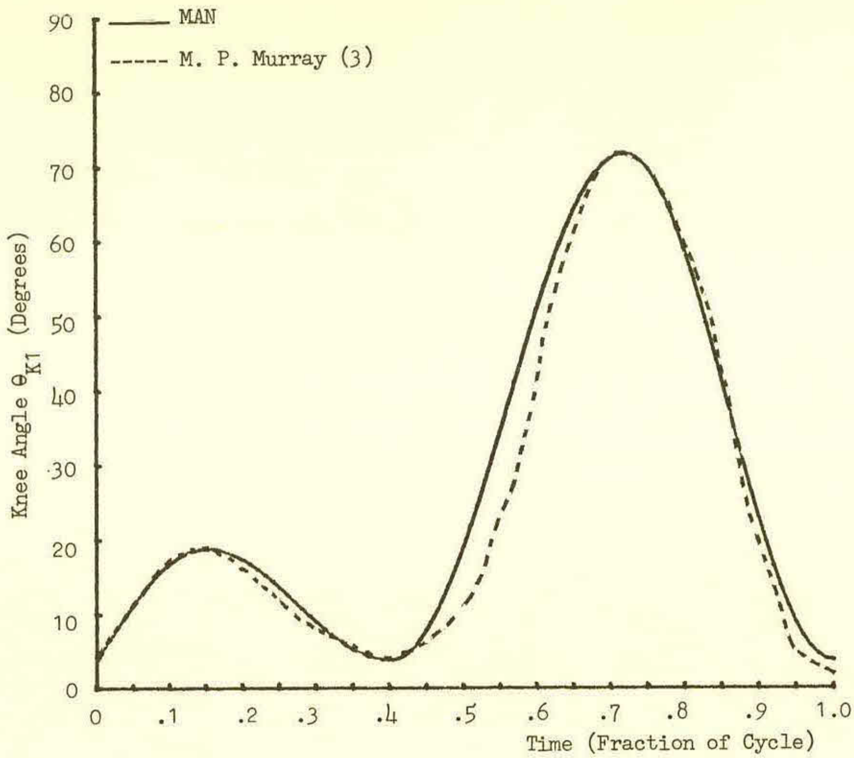


Fig. 10. Knee angle results for initial set of gait parameters.

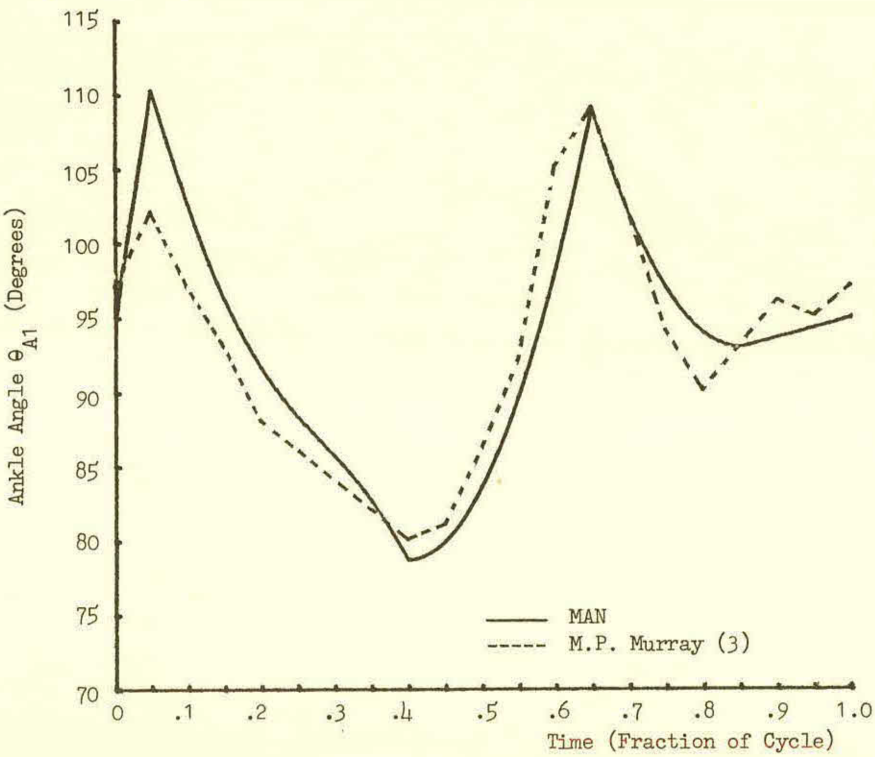


Fig. 11. Ankle angle results for initial set of gait parameters.

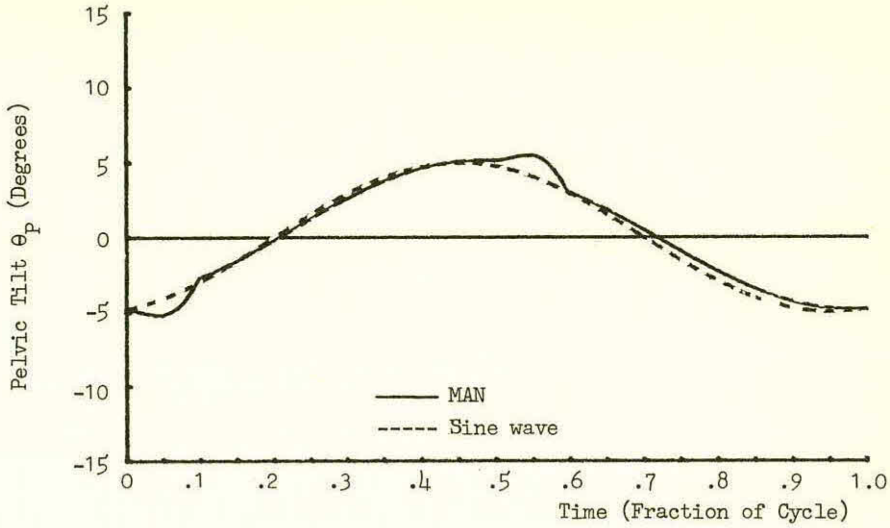


Fig. 12. Pelvic tilt angle results for final set of gait parameters.

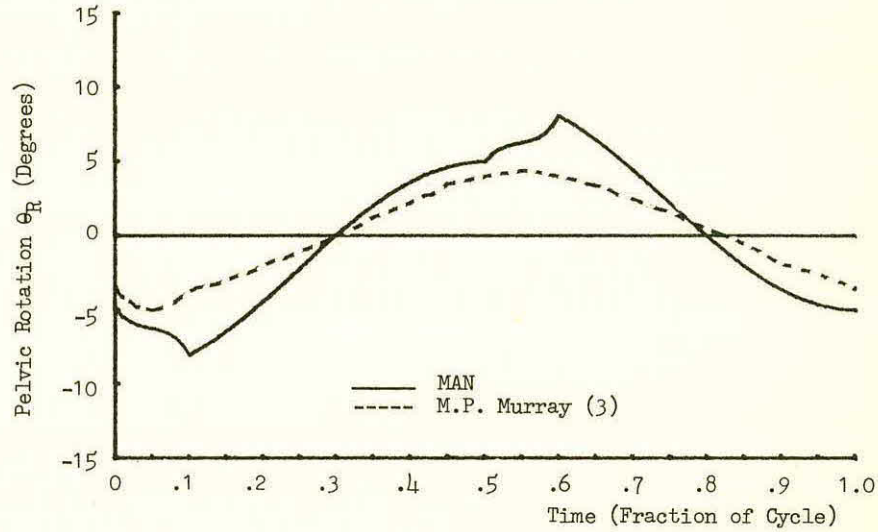


Fig. 13. Pelvic rotation angle results for final set of gait parameters.

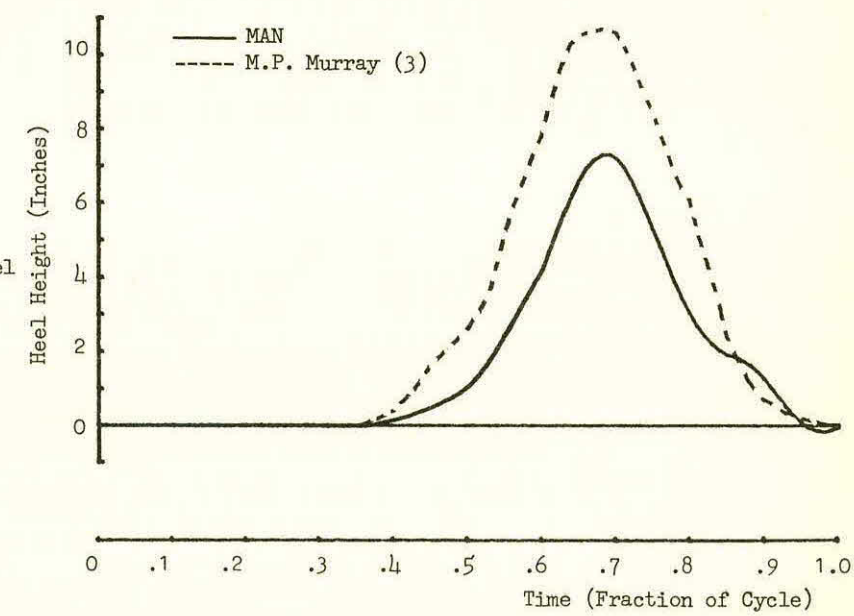


Fig. 14. Vertical displacement of the heel results for final set of gait parameters.

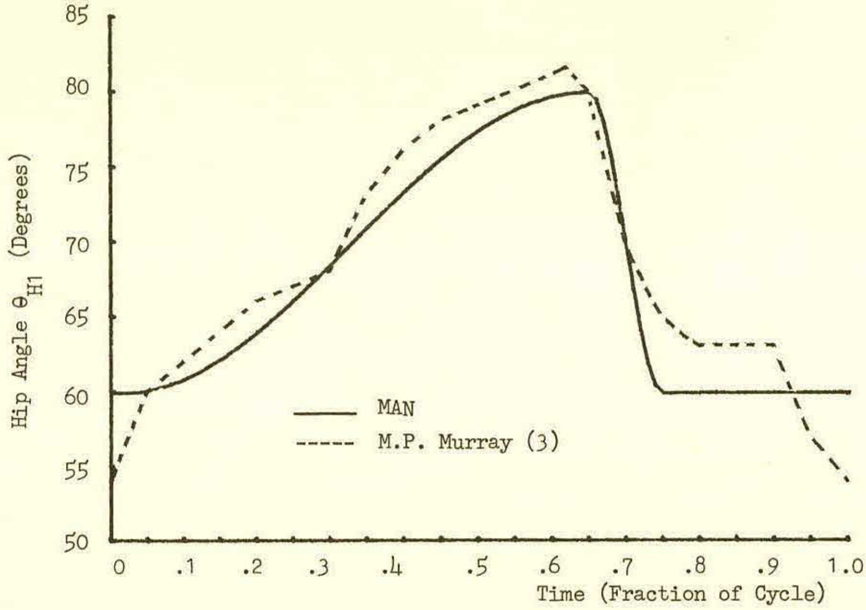


Fig. 15.--Hip angle simulation for paralysis agitans.

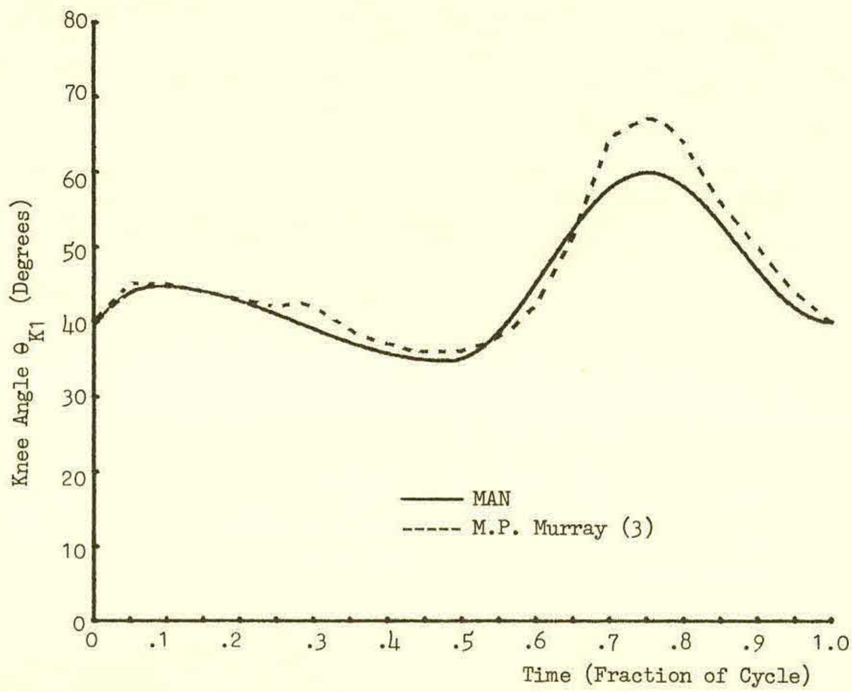


Fig. 16.--Knee angle simulation for paralysis agitans.

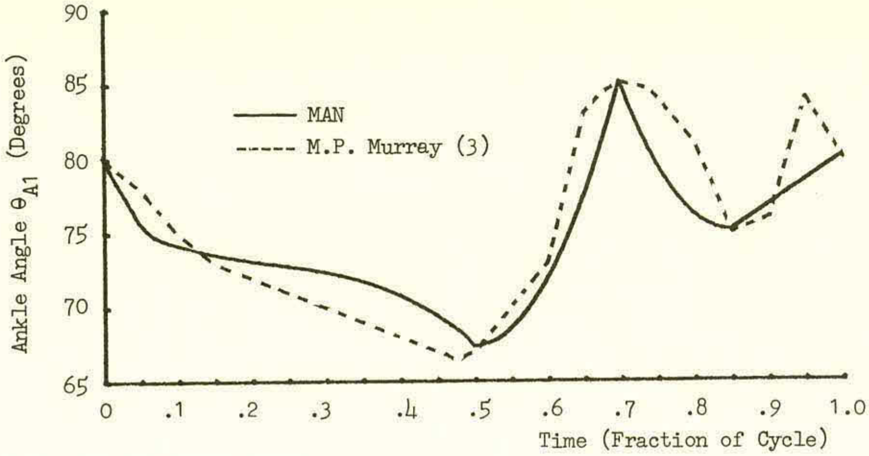


Fig. 17.--Ankle angle simulation for paralysis agitans.

REFERENCES

1 Hartrum, Thomas C., Computer Implementation of a Parametric Model for Biped Locomotion Kinematics, Ph.D. Dissertation, The Ohio State University, Columbus, June, 1973.

2 Saunders, J. B. deC. M., Inman, V. T., and Eberhart, H. D., "The Major Determinants in Normal and Pathological Gait," J. Bone & Joint Surg., Vol. 35-A, No. 3, 1953, pp. 543-558.

3 Murray, M. P., "Gait as a Total Pattern of Movement," Am. J. Phys. Med., Vol. 46, No. 1, 1967, pp. 290-333.

4 Murray, M. P., Drought, A. B., and Kory, R. C., "Walking Patterns of Normal Men," J. Bone & Joint Surg., Vol. 46-A, No. 2, 1964, pp. 335-360.

5 Drillis, R., and Contini, R., Body Segment Parameters, Tech. Rpt. No. 1166.03, Sept., 1966, School of Engr. and Sc., New York Univ., University Heights, N.Y.

6 Beckett, R., and Chang, K., "An Evaluation of the Kinematics of Gait by Minimum Energy," from Biomechanics, Proc. of the First Rock Island Arsenal Biomechanics Symposium, April 5-6, 1967, Bootzin, D., and Muffley, H. C., ed., Plenum Press, New York, 1969.

7 Chow, C. K., and Jacobson, D. H., Studies of Human Locomotion via Optimal Programming, Tech. Rpt. No. 617, Oct., 1970, Div. of Engr. and Appl. Physics, Harvard Univ., Cambridge, Mass.

VOCAL SWITCHING AID FOR QUADRIPLEGICS

by

Robert A. Curtis

Summary. A switching device is described which uses an appropriate timed sequence of audio noise bursts to control electrical relays, and hence, aid a quadriplegic to perform minor tasks. The versatility of the unit is enhanced by the binary counting of the noise bursts and by the ability to add dependent functions to the various output relays.

Introduction

In order to aid a quadriplegic, a relatively simple means was sought to allow the quadriplegic to switch on or off, several different electrical items independently. Another important consideration was to keep the cost of the unit as inexpensive as possible and yet to allow for future additions and more sophistication.

Several items were considered in the design of the unit. The first consideration was the method of switching. For economy and ease of construction, it was decided to use a timed series of audio noise bursts. Second, the quadriplegic should be able to switch on or off any relay independently of the other relays. Third, a fairly large number of relays should be available or be able to be added with a minimal amount of difficulty. In line with having a large number of choices available, a minimal amount of effort should be required to switch on or off any particular relay. A binary timed sequence of audio noise bursts was found to be easy to implement and to easily control a large number of relays in a minimum amount of time.

Operation

The operation of the unit depends on a properly timed sequence of audio noise bursts, which will be referred to as "clicks" in this paper. Each relay is assigned two numbers; one number is for turning that particular relay on, and the second number is for turning it off. The "clicks" are used to count up to a particular assigned relay number. If the sequence of "clicks" are properly timed, that particular relay will be turned on or off independently of the other relays. The sequence of "clicks" is binary in nature and is arranged so that a slight pause between "clicks" will double the value of each "click" counted. The timing of the "clicks" is dependent only on the last "click" so that no overall timing sequence is required. As an example, the number four would be either: "click" - "click" - "click" - "click" or "click" - "click" - pause - "click". If a long pause occurs with no "clicks", the unit will set the relay, using the number just counted. The actual timing is such that the count will double if the space between "clicks" exceeds about .3 seconds and will reset the counter and set the appropriate relay if more than one second elapses with no "clicks".

Block Diagram Description

An overall block diagram is shown in Figure 1. The audio processing is very simple and consists of an audio amplifier and/or rectifier circuit. The

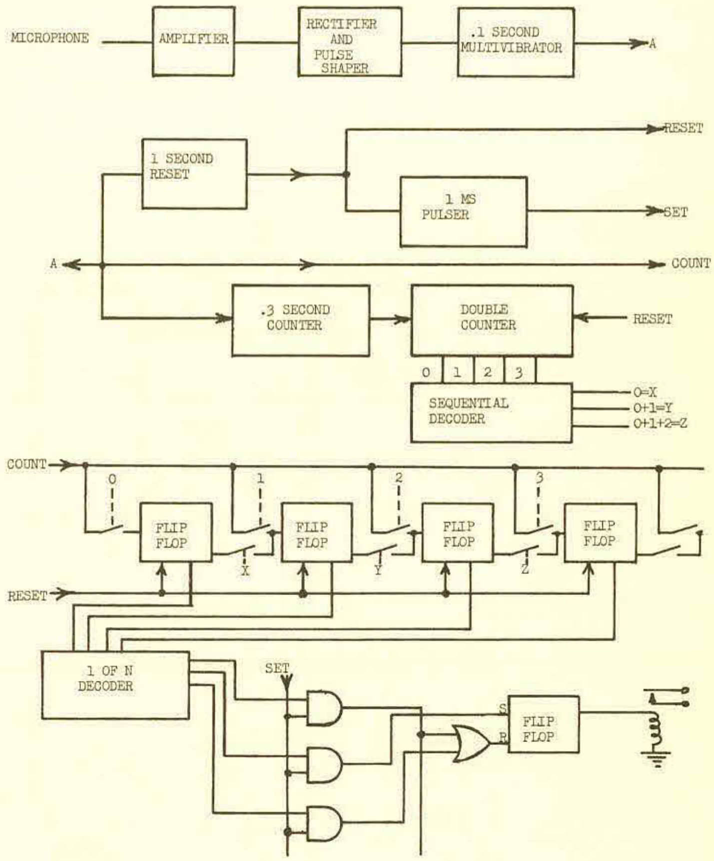


FIGURE 1. Block diagram.

rectifier drives a .1 second multivibrator which prevents multiple pulses in less than .1 second. The next section consists of the timing circuits and the counter drive. The .3 second count detects a pause greater than .3 second between "clicks" and advances the doubling counter. Each advance of the doubling counter doubles the value of each "click".

The counter section consists of a series of flip-flops, FET switches, a one out of sixteen (or greater) decoder, and flip-flops to hold a particular relay position. The FET switches are used to apply the count pulse to the appropriate flip-flop. Each time a pause occurs between "clicks", the value of each "click" is doubled by applying it one step further along the flip-flop counter.

When one second is exceeded with no "clicks", a set pulse turns on (or off) the appropriate output

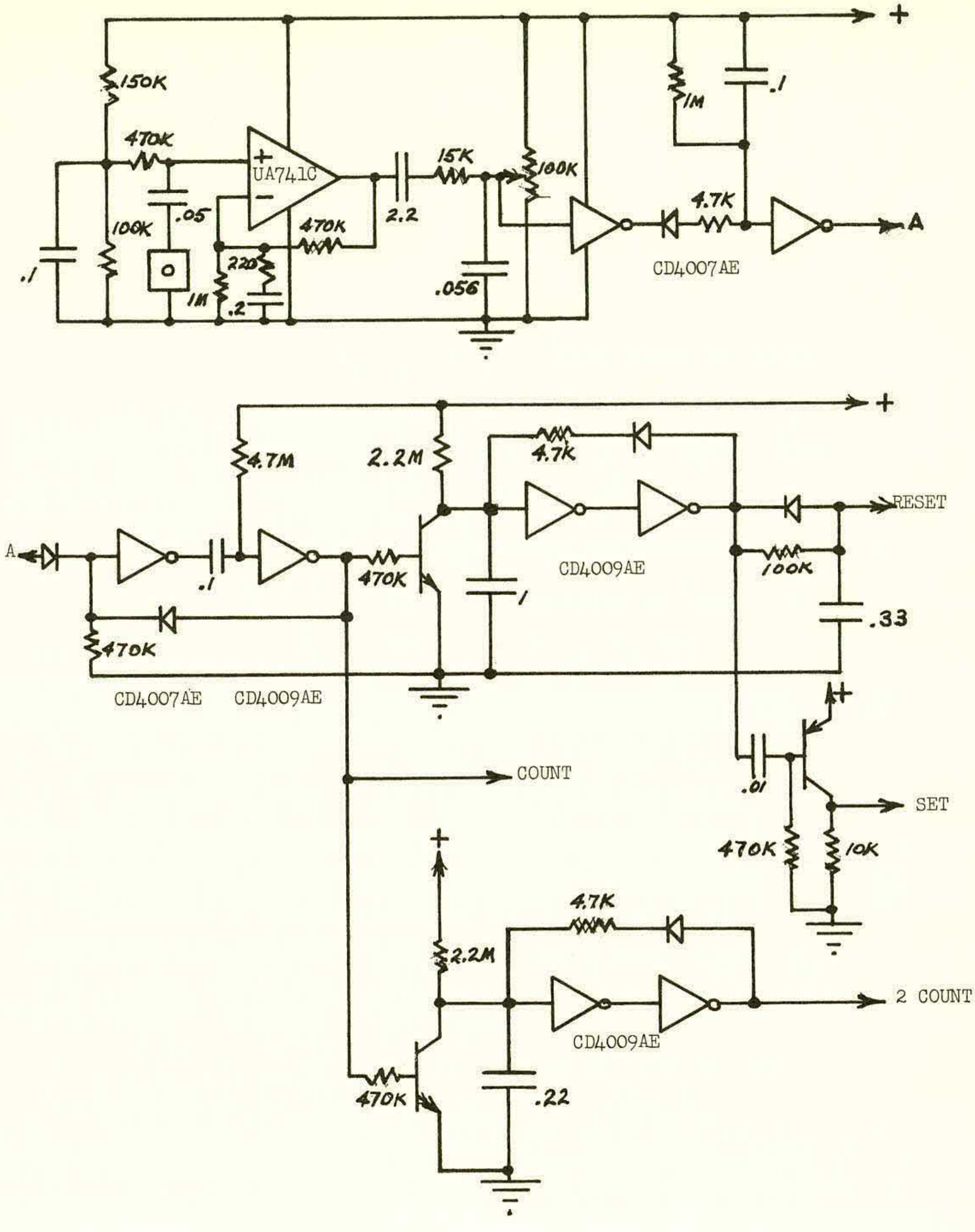


FIGURE 2a. Schematic.

flip-flop and relay. A slight delay is introduced between the reset and the set so that the flip-flops are not reset before the output is switched. In connection with the delay, the flip-flop counter does not see the first "click' so that the one out of sixteen decoder must be connected accordingly. This could be overcome by introducing a slight delay in the "Count" line if necessary. The missing first click wasn't required for the operation of an experimental unit. As the unit was initially configured, one number was set aside that would turn off all the output flip-flops. A complete schematic of the original unit is given in Figure 2. All the logic used is standard COSMOS logic.

Future Suggestions

The switching unit as described is quite useful in aiding a quadriplegic to turn several electrical items on or off. With the addition of another switching unit which is dependent on a particular relay of the above unit being turned on, several ad-

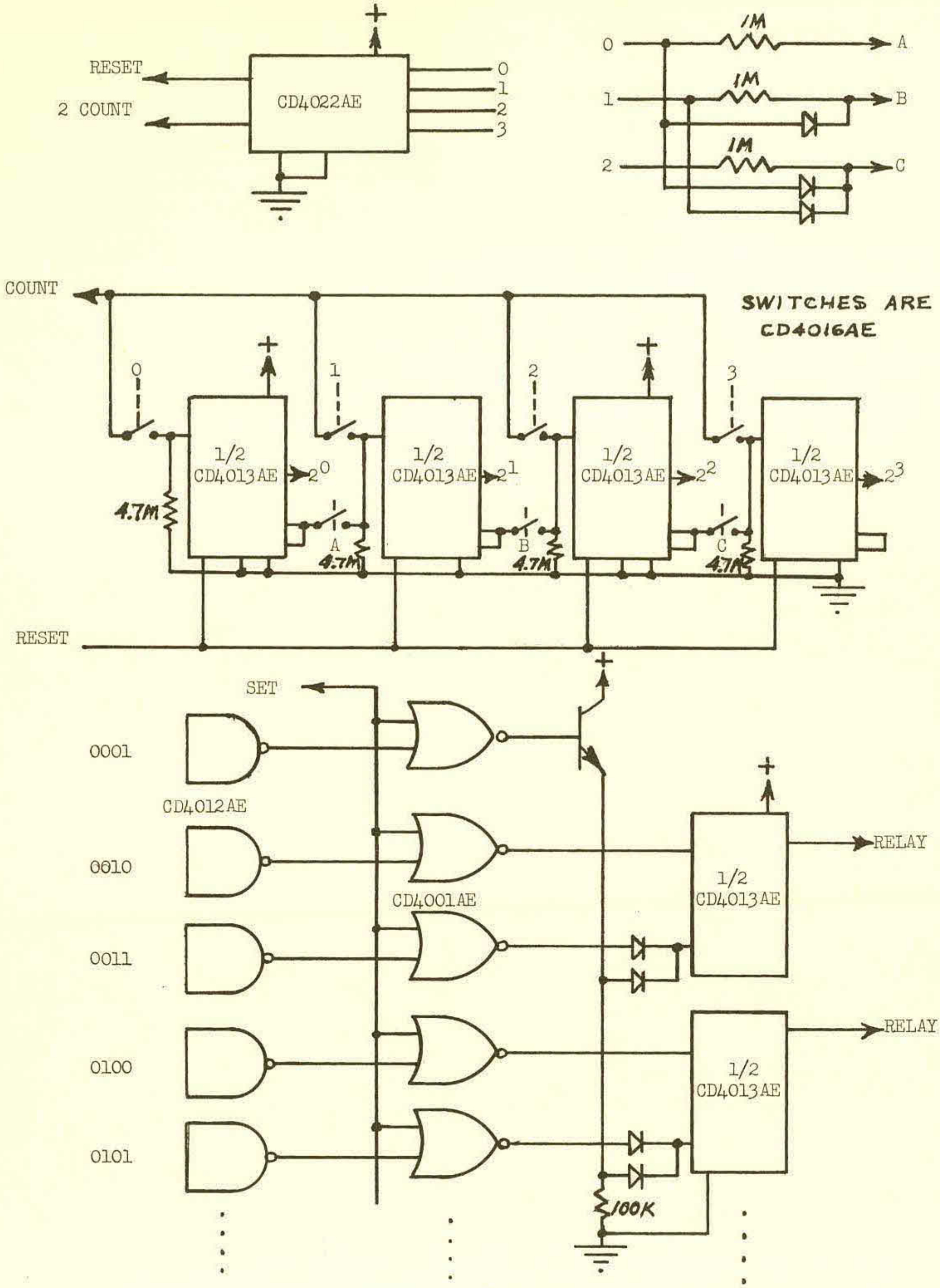


FIGURE 2b. Schematic.

ditional functions could be realized. For instance, a simple high or low frequency detector could be used to adjust a radio or television volume control. A simple addition to the switching unit could be set up to both dial and to answer a telephone. Dialing a telephone would require that a number higher than ten be set aside for the telephone. Once this number is selected, numbers one through ten could be used without triggering the relays normally assigned to the numbers one through ten.

Bed adjustments could also be done with an arrangement such as that used with a telephone plus a high or low frequency detector. It must be remembered, however, that as the unit does work off audio pulses, a random noise in the vicinity may turn on or off some particular relay. The possibility of accidentally triggering both a high number relay and also a dependent function of that relay become rather remote and can probably be used quite safely. Here, however, the safety of the person operating the unit should be considered in light of an accidental relay switching. Good microphone placement and level adjustment would help to minimize false triggering. An experimental unit worked with very few false triggerings when operating either a TV or a radio.

Conclusion

This unit's basic usefulness lies in its relative inexpensiveness and its ability to easily provide a fairly large number of switching functions. A second advantage is the ability to easily add many dependent functions to those already present. While the unit has a disadvantage of false triggerings, its inexpensiveness and many functions offer a quadriplegic considerable aid in performing minor tasks.

MOBILITY AID FOR QUADRIPLEGICS

by

Joachim A. Maass
Vera S. Maass
University of Missouri
Kansas City, Missouri

Summary. A new mobility aid for quadriplegics is proposed which does not require special training for its use. Application, which conserves the user's energy, allows for full visual control of its function, and it is nearly invisible. The device is basically a telemetering system which uses electrical frequency changes corresponding to relative displacement of the jaws as control variables. Passive resonant circuits are worn in the mouth and are more or less detuned by the relative movement of an inductively coupled reactance. The circuits are periodically triggered into oscillations which are subsequently received through an easily concealable one-turn coil around the neck.

Processing circuitry which may be worn in a garment or installed in a wheel chair derives control signals from the frequency measurements which may be used to control the movements of the wheel chair or of other assist devices.

INTRODUCTION

Quadriplegia is the paralysis of all four extremeties. A person in this condition is in desperate need of help with almost every function and he cannot afford to be choosey; he has to accept the aid of whoever is available. The psychological stress resulting from the quadriplegia itself is often compounded by tension due to an incompatible relationship which can develop between the quadriplegic and his helper. The well-meant friendly chatter of a volunteer helper may frustrate the patient because he knows that he cannot escape under his own power. On the other hand, the patient cannot tell the friendly volunteer to shut up. Therefore, if we want to provide the quadriplegic with a mobility aid, we must develop a device which increases the number of options for independent activities available to him.

A wheel chair which merely relieves a helper of some of the physical burden required for moving the quadriplegic apparently is not enough. The wheel chair becomes a mobility aid only if it is powered by a motor and if all motions can be controlled by the quadriplegic. There are also other requirements which must be considered if the gains are not to be offset by the damages:

- The mobility aid should contribute to the user's independence without adding to his burden.

- It must be simple to use, be as versatile as possible, and above all, it must be acceptable to the user.

The existing designs often violate one or more of these requirements, for instance, a device which requires surgical intervention or special skills for the installation of electronic devices, such as myo-electric transducers, is not simple to use and may not be acceptable to the user. Since the quadriplegic is already very sensitive about his condition, contrivances which disfigure the user are just as undesirable as devices which interfere with important sensory functions and which quickly fatigue delicate muscles, such as is the case with most eye-movement detectors. Any device consisting of visible wires hanging from the mouth, nose, or ears, disfigures the user and is not likely to be accepted. Anything that requires the user to behave oddly and conspiciously or requires visible technical appendages is also not acceptable. Although wars and the increasing number of accidents on our highways have made people become more or less accustomed to seeing their fellow man in a wheel chair, they do tend to stare at a person with unusual appendages. Even if none of these drawbacks exists, a device may still be unacceptable if it has "a mind of its own", in other words if its durability is low or if its control functions are unreliable.

METHOD

The device which is introduced in this paper does not require special training for its use or applications. It conserves the user's energy, and allows full visual control of its function and is nearly invisible. The device is basically a simple passive transponder system which telemeters relative changes in the positions of the jaws and uses these to derive control variables for the operation of a motorized wheel chair.

The system relies on the components shown in Figure 1. The user wears several electrical resonant circuits in his mouth. These circuits can easily be incorporated into a cosmetic splint which is attached to the upper teeth, similar to cosmetic teeth coverings used by movie actors to hide blemishes in their natural teeth. A similar device is attached to the lower teeth. It incorporates in the position of the lower central incisors a suitable detuning device such as a piece of ferromagnetic material or a small conductive sheet around the neck. A

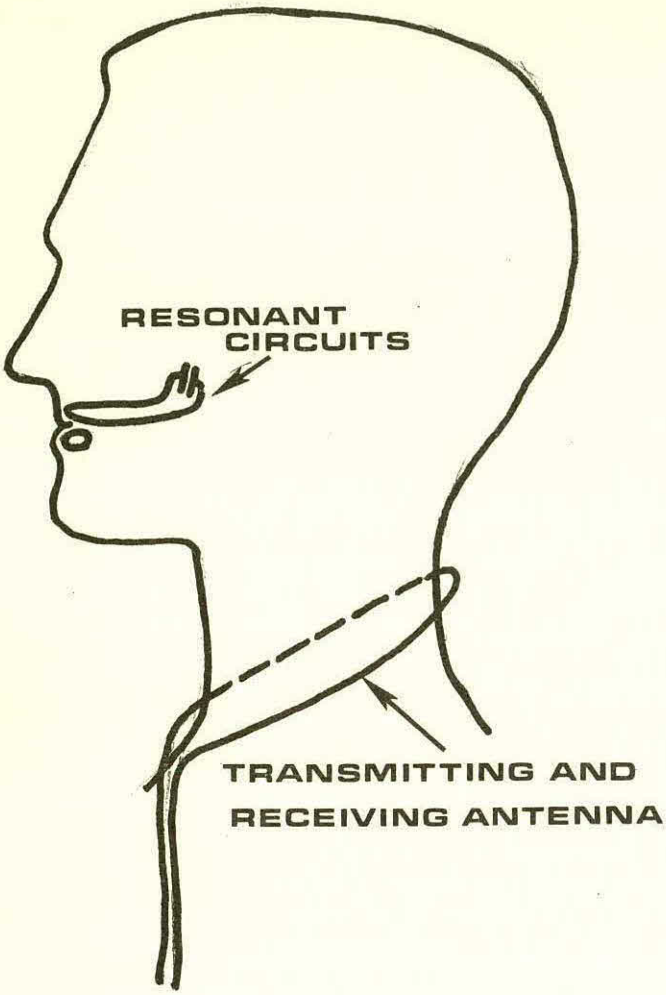


Figure 1. The system components.

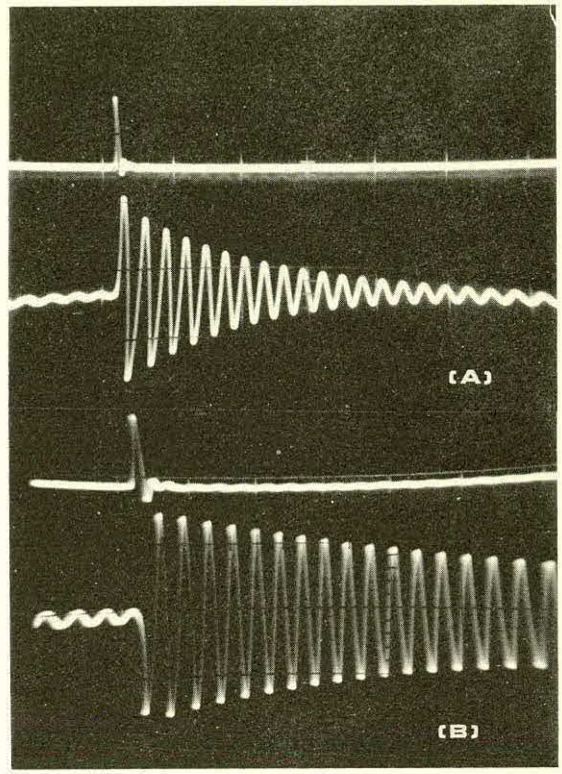


Figure 2. (a) Trigger pulse (b) Triggered oscillation.

one-turn coil, worn around the neck and easily concealed in a shirt collar, acts as a transmitting and receiving antenna inductively coupled to the circuits in the mouth. A current pulse in the coil triggers the circuits in the user's mouth producing oscillations which decay as shown in Figure 2.

Figure 2 shows the trigger pulse (top line) and the triggered oscillation below. By changing the L-C ratio of the resonant circuit the duration of the decay can be altered as shown in the lower part of Figure 2. Again, the trigger pulse and the much slower decaying oscillations are seen. Since it is necessary to obtain control commands for "left", "right", "forward", "reverse" movements and for "no movement", the system must be able to sense corresponding jaw movements. This can be done in various ways, one of which is indicated in Figure 3.

Figure 3 shows the upper dental arch as seen from below. There are three circuits, A, B, and C. The detuning device on the lower teeth can be situated at a point where it will have a greater effect on circuit C than on B or A. Similarly, there is a point where it will

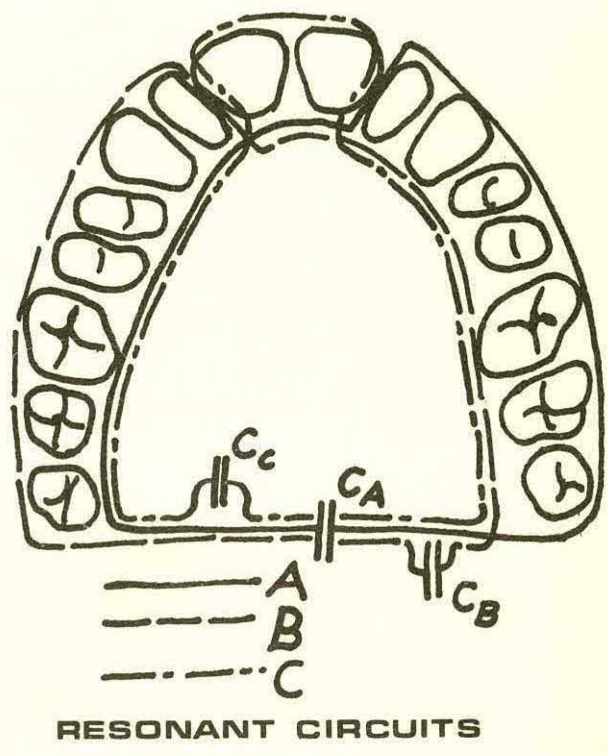


Figure 3. Upper dental arch and the relative positions of circuits A, B and C.

affect circuit A more than circuits B or C. The detuning device then can be moved into a third position where it will have a minimum effect on all three circuits, and, conversely, a maximum effect. This suffices to derive the commands "forward", "reverse", "left", "right", and "stop", whereby the user may decide whether he wants the upper or the lower jaw to be the "direction pointer".

Figure 4 represents a basic diagram of the processor which can be worn in a coat pocket or which could be built into the wheel chair. When the switch below the antenna coil is in the T position, it discharges a capacitor which in turn is charged from a battery, through a resistor. When the switch returns to "R", the receive position, the oscillating signal from one of the resonant circuits is received through the antenna coil which is now tuned by a capacitor, selected by the frequency switch S_f, to the proper frequency. The signal is amplified in A and processed in P. The three resonant circuits are interrogated in a sequence determined by a clock in P. The results are the commands for the four directions and "stop" and also for any intermediate situation such as "go slowly right forward".

The circuit will also detect any electromagnetic interference through the interference sensor, IS. An excessive signal from here will turn off all functions and sound or flash an alarm. However, it is extremely unlikely that an extraneous field could reach such proportions that its magnetic component will induce a voltage in the antenna coil high enough to interfere with proper circuit operation. Furthermore, it is possible to use "receiver gating" synchronously following the trigger pulse and a decay detector for the received signal which would distinguish it sufficiently from interference. All "switches" in the diagram are semiconductor circuits.

It would not require much additional electronics to detect opening of the mouth as in speaking or yawning. One could incorporate a memory into the circuit which ignores such events and holds the last received command until proper operation is restored. The mouth-opening detector could also be used to detect certain patterns in rapidly repeated mouth openings and closings which could serve as turn-on and turn-off signals for the equipment.

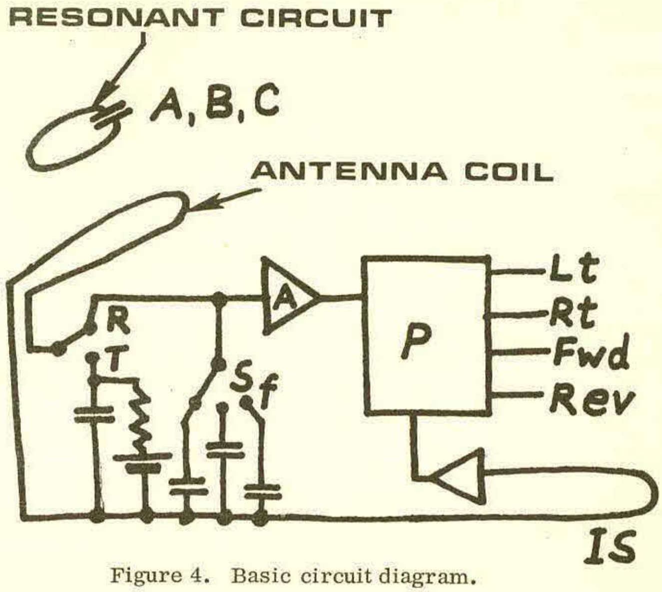


Figure 4. Basic circuit diagram.

The device as shown is a product of an experimental study to prove feasibility. Its basic components have been operated with resonance frequencies of approximately 20 MHz. There is no measurable attenuation of transmit or receive signals when the resonance circuits are inserted into the mouth, neither have saliva or movements of the tongue any effect. The frequency stability is high due to the thermostatic environment of the oral cavity. A trigger capacitor which is charged to approximately 5 volts produces a received voltage of approximately 50 millivolts peak-to-peak for the first cycle. The resonant circuits consist of one-turn coils with appropriate capacitors. The rear portion of the coil curves upward to conform with the shape of the palate.

The device described above is intended to relieve the burden of the paraplegic, to make his life more tolerable, and, above all, to restore his sense of dignity by giving him the ability to function independently.

MOBILITY OF THE BLIND

by

Howard Freiberger
Veterans Administration
252 Seventh Ave.
New York, New York

Summary. In this paper I shall try to give an overview of the subject of Mobility of the Blind. Blind people sometimes are reputed to be able to get around without using any external aids at all. It is felt, however, that a totally blind person really is unable gracefully, safely, and independently to get around in an area outside of the confines of familiar surroundings without employing some mobility means. It is true that a blind person trained in basic mobility skills, often the first things taught when one gets mobility training, is able without any visible external aids to get around after a fashion in limited circumstances. For example, by the use of the forearm position, which is a way of holding the forearm up at about face level, palm forward and across the body in such a way as to detect obstacles which might bump the body, a blind person can negotiate with some safety in some surroundings. He will also often trail with one of his hands along an adjacent wall. This means of mobility, however, not employing any external aid as such, cannot be considered suitable for extended or frequent travel in unfamiliar locations, and is not really a graceful way of traveling.

The Sighted Guide

A very widely used travel means is the sighted guide. Use of the sighted guide is a learned skill when done at its best, and involves more than walking arm in arm with a sighted person. Rather, there is a particular grip, where the blind person grips the elbow of the sighted guide between his thumb and fingers, and where he stands a bit behind the sighted guide so as to narrow the front of the approaching pair enabling better negotiation of restricted passageways and doorways. The blind person can feel the movements of his sighted guide and can sometimes sense some of the things the sighted guide sees. Thus the elbow of the guide in a specially construed sense becomes an anatomical element in the visual system of the blind person.

The Long Cane

A rather commonly used external mobility means for the blind is the long cane or typhlocane, or Hoover cane (1). This device, probably a follow-on from earlier attempts at using staffs and walking canes in the past, was developed essentially into its present form in the later years of World War II and the early post World War II years within the Army and later the Veterans Administration system for blind rehabilitation. The cane is much longer than an ordinary orthopedic cane, measured roughly from the tip of the sternum to the side of the foot extended. The technique of using the cane is to have it centered on the body and projecting forward and downward, and then swept from side to side in an arc about the width of the body, touching the ground at the extremes of the arc such that the cane tip touches where the foot will next fall. This arrangement enables the blind person to assure himself of the presence of the ground, and also gives him information as to the nature of the ground where his foot will next fall. It can be seen that this provides for relatively safe passage. If there is something in the way in the region traversed by the swinging cane, collision with such object usually signals its presence to the blind person sufficiently early that he can avoid body contact. Also the touching of the ground at the lateral extremes of the arc enables the person to detect the occurrence of a curb or drop-off in sufficient time to prevent stumbling or falling at the curb. The cane tip touching the ground provides information about the nature of the ground, and the shaft of the cane touching objects which may intervene also gives information as to their location and sometimes their identity.

A typical long cane which has been standard for some years now within the Veterans Administration (2) has a small crook at the top end. This crook can be used to hang the cane on a door or hook when it isn't needed. People have even been known to hook the cane in the back of their collar when traversing a cafeteria line requiring the use of both hands to guide the tray along the line. The grip of the cane is made of a modified golf club grip usually of some rubber-like material with a flat side to help in orientation of the cane. The grip also serves as a rest for the index finger which points down the shaft creating the feeling that the cane itself is an extension of the index finger; sort of a natural probing and sensing device.

The hollow shaft of the cane is generally made of an aluminum alloy and sometimes is covered with a special tape to provide light reflecting capabilities at night. The lower portion is often colored red in accordance with practice relating to canes used by blind people, and the tip itself in the VA's standard is generally made of nylon. Some persons seem to prefer furniture glide tips but VA experience indicates that nylon provides a good compromise between a material which minimizes catching and sticking in the ground, yet which gathers some information about the ground it is touching.

A long-cane user properly trained, it is felt, can successfully negotiate almost any travel route with a reasonable degree of safety, grace, and speed. When using a long cane, a blind person naturally should be fully aware of his surroundings and must use all cues to compliment the cane and achieve a good mobility performance. He must use his ears, sense of smell, his kinesthetic sense, and the tactile sense, including feeling warmth on the skin or the passage of air across the face. A composite system is thus achieved which provides reasonable mobility capabilities for the blind person.

The Dog Guide

Another well known and reasonably widely used system of mobility for blind people is the dog guide. The term dog guide written with small letters is used as the generic term referring to a number of special dogs trained by organizations such as The Seeing Eye, Guide Dogs for the Blind, Pilot Dogs, or Leader Dogs for the Blind (3). The dog is a mobile, intelligent, trainable, hard working, and friendly kind of mobility aid and can do a remarkably fine job in many circumstances. The dogs generally are specially bred and selected in their early youth, given special training at the training center, and then after reaching a certain age are brought together with their prospective blind master at the training center. Both are given additional training so that the pair can perform as a team. The dog is a working dog, and at best performs when there is a close relationship with his master. Sometimes this close relationship poses problems and the user of the dog must learn to overcome such things. The dog being a living creature has to be fed, aired, and walked, and sometimes he becomes ill. People with dog phobias obviously are not good candidates for dog-guide training. A small percentage of blind people are in fact dog users, perhaps only two percent or less, but those who do use the dog and use him well certainly find him to be an excellent means to aid in their achieving independent mobility.

Electronic Mobility Aids

Probably as early as photosensitive materials became known and electrical circuitry could be used as adjuncts to them, inventors have thought of using such combinations as mobility aids for the blind. We know that as early as 1912, Doctor E. E. Fournier d'Albe developed his optophone initially as an environmental sensing device (4). Through the years since then various attempts have been made to use electronic means to provide a mobility aid for the blind. During the later years of World War II, when the public was well aware of the environmental sensing done by radar and sonar, inventors renewed their efforts to apply similar principles in developing a mobility device for the blind. By 1946, Lawrence Cranberg, who was then working at the Signal Corps Laboratories in New Jersey, developed his mobility device (5). This unit was hand-held and about the size of a large camera. It employed an incandescent light source which was interrupted by a chopper disk. The light source was focussed by a lens to illumine a small region ahead of the device. Energy pulses would reflect back from an object ahead into the receiving optics. Photocells would sense the pulses and convey the information to the user. The device worked on the principle of optical triangulation. Range information was recovered by passing the return light through a chopper wheel having different numbers of holes at different radii. Depending on the triangulation, the return beam would thus be interrupted at a frequency depending on the range to the reflecting object. The information was conveyed to the blind person by vibration of the handle of the device according to the frequency of interruption of the return energy.

The user thus learned of the presence of obstacles and of their range by feeling vibrations through the handle of the system. He knew the direction or azimuth by knowing which way he was pointing the device. The Signal Corps device was constructed by RCA in about 25 copies, but it was found to be somewhat heavy to carry and not productive of the kinds and quantities of information that would compensate for the trouble of using it, thus it never became widely accepted by blind people. A few years later a precursor firm to Bionic Instruments, Inc. tried to construct a similar device using the somewhat later technology then available. This resulted in the Model G-5 Obstacle Detector which was somewhat lighter, worked without mechanical chopper disks and used an improved source of illumination. This device was bulkier and heavier than one would have liked for the kind of information it returned, and it also did not receive widespread use.

In the early fifties, a special electronically equipped cane was developed at the Franklin Institute in Philadelphia. This cane had the capability of sensing the capacitance to the ground from its tip. The hope was that this cane, when approaching a curb for example, would give early warning by an audible signal activated by the change of capacitance. This would enable the person to handle the situation with a little more lead time than with an ordinary cane. The early warning provided by the system was so slight that this cane too never found its way into routine use by blind people.

The inherently excellent characteristics of the long cane as a mobility aid, and the promise that environmental sensing at a distance should be able to contribute to the mobility performance, led in 1962 to the thought that electronic sensing means should be built into a cane. The engineers at Bionic Instruments, Incorporated, decided to mount an environmental sensing system in a cane. The concept was to improve upon the ordinary typhlocane. This was accomplished by having the new cane "look" out in three directions; upward to "clear" the region through which the head would pass, straight-out ahead to detect obstacles beyond the reach of the cane tip which might come into the traveler's path, and down at the terrain to detect discontinuities such as curbs and drop-offs. It should be noted that the ordinary cane does not provide safety from collision with objects at head level such as boughs of trees or structures projecting from walls in passageways. While the ordinary cane does detect things ahead, it can do this only by touching them, whereas if sensing could be accomplished at a distance, a person might be able gracefully to avoid striking the object rather than coming into cane contact with it. The ordinary cane also can sense discontinuities in the terrain, but the time then remaining for the blind person to react is quite short. By having an electronic system "looking" several feet ahead of the cane tip, additional early warning can be obtained which might enable the person to travel more comfortably and more securely.

Passing through several models and extending over a period of years, the laser cane finally came into being in its most current form in 1973 (6). It is not much grosser in any of its dimensions than an ordinary cane, being only a bit thicker and somewhat heavier to accommodate the electronics and batteries. The three-channel idea is preserved in the current version. One channel looks up at the zone through which the head would pass as the traveler moves forward. If there is something there, reflected energy signals the cane which produces a high-pitched beeping tone which the user hears. If there is something out ahead, the cane also senses this. It provides an indication by giving a tactile stimulation against the finger held along the flat of the cane and also optionally by giving a tone of intermediate pitch.

The third channel directs a beam down toward the ground, reaching the ground several feet ahead of the cane tip. If the ground is there and adequate energy is reflected back into the receiving optics, the cane gives no signal indicating all is clear as far as the terrain is concerned. If, however, something disturbs this beam, such as a curb or drop-off, the user will be signaled by a low-pitched beeping tone, indicating some discontinuity in the terrain or at least in the terrain-sensing beam path.

Analyses of actual user experiences with a similar earlier cane of this type, built in ten copies and employed in an evaluation program involving eight blinded veterans, have yet to be published. Indications are, however, that for some persons the added information provided by the capability to sense at a distance may, with practice, contribute to a less stressful and more graceful mobility performance.

Devices employing sound energy to probe the environment have also been considered as mobility aids for the blind. Several years ago, Professor Leslie Kay of Birmingham, England, now of Christchurch, New Zealand, developed his hand-held torch for the blind. This device looks somewhat like a large flashlight and has two transducers mounted into its forward portion, one to transmit ultrasonic energy into the environment, the other to receive echoes from objects in its field. The user is apprised of the presence of such objects by having a tone presented in an earphone. The device uses swept frequency ultrasonic energy to insonify the environment and then compares echoes with the transmitted sound to produce a difference-tone audio indication that something is out ahead. The greater the range to the object detected, the longer the time of flight of the sound and its echo and hence the greater the difference in frequency between the transmitted and received energy and hence the higher the pitch of the auditory signal heard by the user.

A hand held torch such as this one proved not to be the best configuration for the device, and the inventor next worked to produce a pair of spectacle mounted devices, now called the Binaural Sensory Aid. The Wormald Vigilant Company of Christchurch, New Zealand has built a pilot run of these devices which have been used in several countries throughout the world for evaluation purposes (7). At the present time, in response to feedback from the evaluations and a desire to incorporate some design changes, the company is building a new device, giving promise of being an improvement over the first model. The new unit is expected to become available in mid-1974. The evaluations conducted to date indicate that for some people, after they have adequate training, this aid provides useful additional information about the environment to enhance mobility. It is not believed that the Binaural Sensory Aid alone can serve as a general mobility device for all circumstances. At the present time it can be used best in conjunction with a long cane or possibly with a dog to give an overall mobility system considered safe and effective.

Another aid to mobility from New Zealand is the Mowat Sonar Sensor. This is a small device about the size of a package of king-sized cigarettes, intended to be a secondary mobility aid, kept in the pocket most of the time. When it is needed, it is taken out and used to scan the close-by environment. It could be helpful in locating doorways, various poles in the street, bus stops, telephone booths, spaces between parked cars, and so forth. It works by insonifying the environment out ahead with ultrasonic energy and receiving echoes from things which may be in its field. Indication of something out ahead is given by the whole device vibrating in the user's hand. It has an armature inside which when activated upon by receipt of an echo will cause the whole case to vibrate and thus transmit the information. Some indication of range is given as the vibration rate of the case is inversely related to the range. An object or structure 20 feet away will cause vibration at about 16 Hertz, and as the distance decreases to about one foot the frequency goes up to about 56 Hertz. Four of these units have been purchased by the Veterans Administration and are currently being evaluated, one each at the three blind rehabilitation centers and one at the central sensory aids laboratory. Results are not yet available to enable one to say just what the value of this instrument will be.

Another mobility aid using ultrasonic energy is the Lindsay Russell Pathsounder developed by Lindsay Russell in conjunction with the Sensory Aids Evaluation and Development Center at MIT. The pathsounder is worn on a neck strap and hangs down over the chest, looks somewhat like a large camera with two horns on it, these for projecting and receiving the sound upon which the sensing-at-a-distance system depends. If there is nothing out ahead of the unit, it remains silent. If an object comes into its field, at a distance, it first signals the user by a series of clicks which are heard from two small speakers mounted in the neck strap close to the user's ears. As the user comes closer to the object, at about three feet, the rate at which the clicks are heard increases noticeably indicating that the person is quite close to the object. The device seems to have promise as a mobility adjunct since it is worn on the chest, leaving both hands free, does not impede the use of another mobility aid such as a cane or a dog, and warns with a simple signal of objects out ahead beyond the reach of a cane. At the present time the unit is being physically improved to make it somewhat smaller and lighter and to make the horns, to which some people objected for cosmetic reasons, a little less noticeable.

Another recent development in the general area of mobility aids for the blind is the eyeglass-mounted Mims Seeing Aid. This device is actually a pair of spectacles.

Two brass tubes containing electronic and optical parts, on one side a transmitting tube, on the other a receiving tube, are mounted into the temple frame. The tubes contain battery power sources, a light-emitting diode (LED) which sends out a beam to probe the environment, and a photosensitive receiving cell. Receipt of reflected LED energy produces a buzzing signal which is delivered to the user through a small tube leading to the ear. Thus this small, light, head-mounted unit warns the person at a distance of the presence of something in the area being illuminated. It does seem to have promise as an adjunct mobility aid and is currently being evaluated for this purpose.

Audible and Haptic Maps for the Blind

Another kind of mobility aid for the blind which serves the orientation function of mobility is of course the map. Maps usable by blind people can be either in the audible form or in the haptic form. In the paper by Bruce Blasch at this Conference, the audible map will be discussed in detail; so I shall not consider it further. The haptic map, in use for quite some time in educating the blind, has recently been developed in a new form to be used as a mobility tool in certain locations (8). One of the first examples of this method is a fairly large map of the MIT campus made on a plastic material about a quarter of an inch thick containing information on both sides of the sheet. On one side is a plan of the campus with buildings and roadways, the river, and many other bits of information indicated in relief. There is also a special symbology used to mark the map in relief and to explain the features of the area. On the back of this map, in corresponding locations just beneath the spot on the front where an object occurs, there are further explanations as to what is at that point. Special symbology and braille notations are used on the back to explain details further. Maps of this kind are also currently being evaluated to determine their usefulness and value as mobility adjuncts for blind persons.

Conclusion

From this brief survey of mobility for the blind, one can see that there are a number of methods which can aid the blind person to get around. Paramount among them of course, are those methods which make use of the remaining senses of the blind person, the sense of hearing, sense of touch, smell, and kinesthesis, and those which make use of someone else's eyes, such as the dog guide or the sighted guide. When using any method, of course, the blind person must be alert to interpreting all of the cues provided by the environment so as to insure a safe, graceful, and least stressful progression through space for himself and if accompanied by someone for the guide, too.

References

(1) Subcommittee on Sensory Aids, Committee on Prosthetics Research and Development, NRC, "The Cane as a Mobility Aid for the Blind," 30 pp., 1972.

(2) Veterans Administration, "Specifications for the Long-Cane (Typhlocane)," 8 pp., Feb. 1965 [also included as Appendix I of "Proceedings of the Rotterdam Mobility Research Conference," pp. 243-251, AFB, May 1965].

(3) American Foundation for the Blind, "Directory of Agencies Serving the Visually Handicapped in the United States," 17th Ed., 364 pp., 1971.

(4) Fournier d'Albe, E.E., in Physikalische Zeitschrift, Vol. 13, pp. 942-943, Oct. 1, 1912.

(5) Cranberg, Lawrence, "Sensory Aid for the Blind," Electronics, Vol. 19, pp. 116, 117 and 119, March 1946.

(6) Benjamin, J.M., Jr., N.A. Ali, and A.F. Schepis, "A New Laser Cane for the Blind," Paper 5.2, 4th Annual Meeting, Biomedical Engineering Society, Los Angeles, 29-30 Jan. 1973.

(7) Kay, Leslie, "Sonic Glasses for the Blind - A Progress Report," AFB Research Bulletin, American Foundation for the Blind, No. 25, pp. 25-58, Jan. 1973.

(8) Kidwell, Ann Middleton, and Peter Swartz Greer, "The Environmental Perceptions of Blind Persons and Their Haptic Representation," The New Outlook for the Blind, Vol. 66, No. 8, pp. 256-276, Oct. 1972.

A THIRD GENERATION SELF-TIMING PAGE MARKER

by
Deane B. Blazie
US Army Human Engineering Laboratory
Aberdeen Proving Ground, Md. 21005

Summary. The history of page marking is briefly discussed from the first L-C oscillator page marker through the first self-timing page marker to the third generation page marker using CMOS integrated circuitry. The third generation page marker, described in detail in this report, provides the following improvements over previous devices.

1. Higher reliability
2. Less distortion
3. Lower switching transients
4. Lower cost

It is hoped that this device will encourage those recording printed matter for the blind to use some form of indexing or page marking.

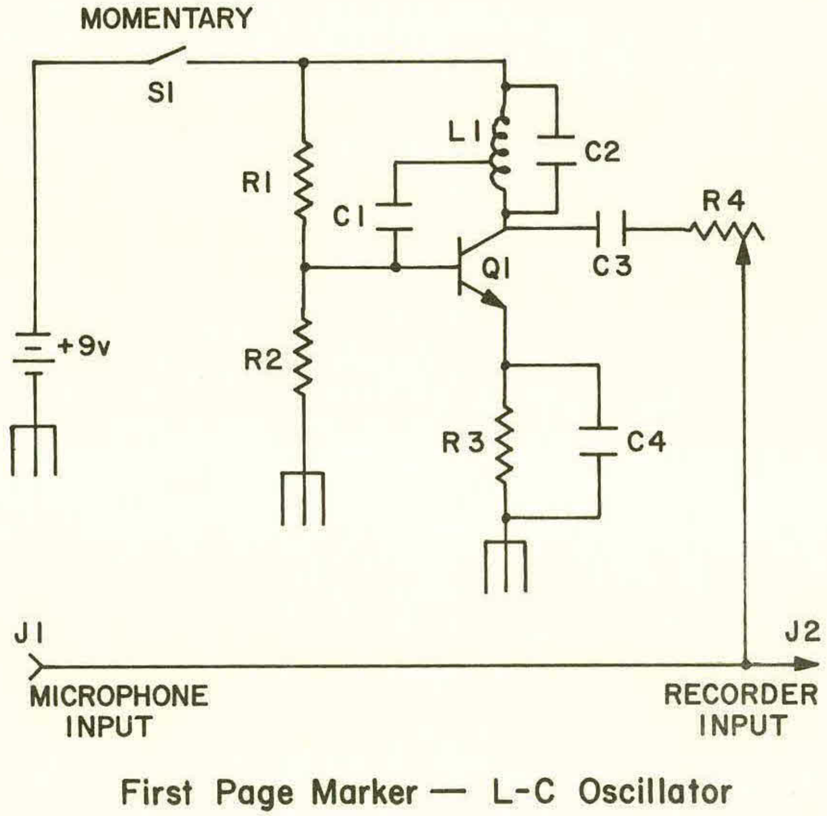
Page marking came into need when magnetic tape recorders first came into the hands of the blind. As the price of magnetic tape recorders dropped and the size became more practical, more and more of the blind community, aware of the advantages of these recorders, began using them.

Students found them useful in the classroom. Blind housewives found them useful for shopping lists and other note taking. Soon magnetic tape recorders were a necessity for blind individuals.

But even before this widespread use, some of the cumbersome aspects became apparent. Unlike most printed reading matter having location markers (e.g. page numbers), (for the blind) a magnetic tape has only a beginning and an end. There are no marks to distinguish one section from another. This means that the listener has to search for a particular passage. Surely, the listener may know the approximate location of the passage, but even dividing a 2-hour tape into 8 sectors means that he must search a 15-minute sector for perhaps a 1-minute passage. Consider further a classroom situation where a week's worth of notes are taken on a single tape. Before each class, a tape must be mounted, threaded (in the case of open reel recorders) and the end of the previous session found in order for the new session to follow it. Well, it didn't take long for blind users to devise techniques to index or "mark" tape passages. As far back as the early 1950's, blind users were recording thumps, pecks and low frequency noises that during fast forward or rewind would catch their attention and mark recorded passages. One ingenious user even cultivated a low frequency flutter of the lower lip to mark passages. This flutter was not as objectionable during playback as a thump on the microphone but during rewind it produced a distinct mark. Soon it was discovered that an open microphone jack, when touched by the hand, introduced 60 Hz hum into the tape, providing a mark. This method gave a mark which was only moderately audible during playback but produced a clearly audible tone when the recorder was in the fast-forward or rewind mode. A 10 to 1 increase in tape speed in these modes produces a 600 Hz tone which becomes audible when the playback electronics are allowed to operate while the machine is rewinding or running fast-forward.

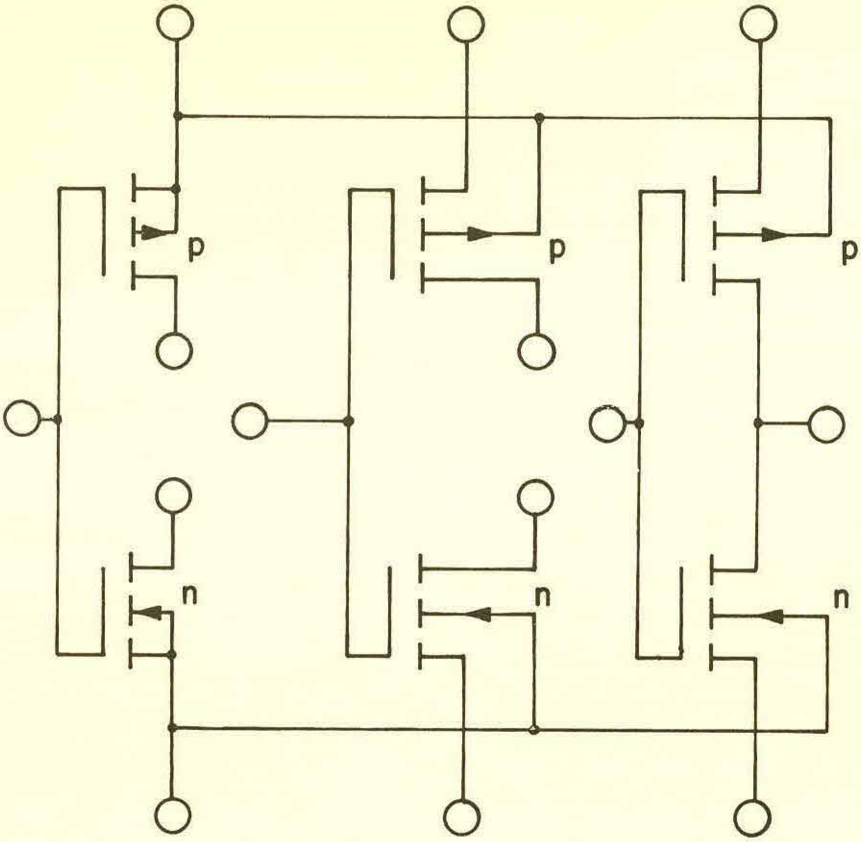
This technique of recording a low frequency sinusoid and playing it back at 10 times the speed to mark a tape led to the development of the first page marker in 1970.

My first marker was designed in conjunction with the Kentucky Department of Education for use by volunteers making recordings for the blind. It consisted of a single L-C oscillator with a push button and level control potentiometer



First Page Marker — L-C Oscillator

Figure 1



4007
Dual Complementary Pair Plus Inverter

Figure 2

enclosed in a mini-box with appropriate connectors. The oscillator produced a 50 Hz tone when the button was held down. A schematic diagram is shown in Figure 1.

While simple and relatively inexpensive, after one year's use, some undesirable features became apparent. First, the circuit drew approximately 25 ma from a 9 volt transistor radio battery. This caused short battery life and unreliability since it was not obvious to the user when the battery was bad. Second, the waveform produced was relatively rich in harmonics which made it audible when the source material was being read by the blind listener. While this is not a serious defect, it was undesirable and eliminated.

Finally, the most serious defect of the page marker was that the volunteers required training to hold the button down for a prescribed period of time. While seeming simple, this procedure required the attention of the volunteer while reading the source material being recorded. We found that even with training, large variations occurred in the duration of the mark. At a record speed of 1-7/8 ips (speed of most cassette markers used by Department of Education) and a fast-forward and rewind speed of greater than ten times this, a one second recorded tone will produce a .1 sec mark which may not catch the attention of the listener.

The latter problem led to the idea of designing a self-timing page marker. Such a page marker would begin oscillation when the button is pushed, remain on for a prescribed length of time and then turn itself off, no matter how long the button was held. It needed to be foolproof in that if the button was held down, only the single mark would be produced. Similarly, if the button is only momentarily pushed, the oscillation will remain on until the time period has ended.

Conferences with the Department of Education, Division of Services for the Blind, and its users led to the finding that a tone duration between 3 and 5 seconds was optimum. Further, a tone frequency between 30 Hz and 50 Hz was agreed upon for the following reason. Frequencies below 30 Hz, while acceptable to the user, became problematic when tapes were duplicated in that when the master tape is recorded, a +10 db signal is required at 20 Hz due to rolloff of the electronics to achieve a level of OVU, but when the tape is duplicated, the 20 Hz tone is reproduced at OVU and is duplicated at 10 db below OVU. Tones above 50 Hz began to be audible at normal reading speed and therefore were undesirable.

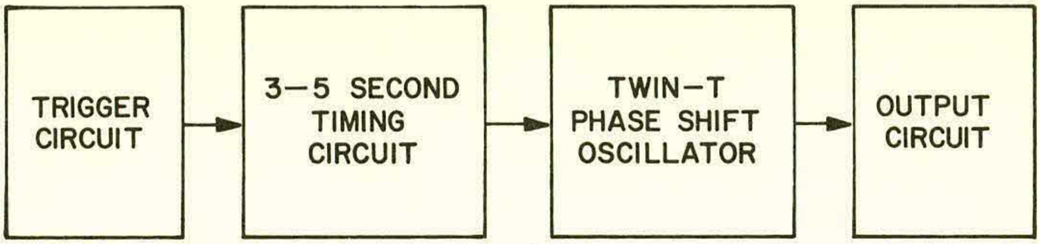
Other considerations in the design of the self-timing page marker were:

1. Output waveform
2. Switching transients
3. Battery life
4. Reliability

The output waveform affects audibility at normal playback levels. A pure (less than 1% distortion) 40 Hz tone played through the machines used by the Department of Education cannot be heard when the recorder is set for normal listening level. This is due to poor response of the speakers, rolloff of the playback electronics and relative insensitivity of the human ear at this frequency. However, unless the distortion is low, the overtones will be audible.

The recorders used by the Kentucky Department of Education have automatic gain control (AGC) with a time constant of approximately 10 seconds. When the page marker is operated, it must not disturb the AGC or else the recorded material will have a reduced level for some time after the mark until the AGC recovers. Switching an oscillator on and off can cause such transients unless care is taken in the design.

The self-timing page marker should draw as little current as possible to conserve battery life and hence be as reliable as possible with battery changes on a yearly basis.



BLOCK DIAGRAM

Third Generation — Self Timing Page Marker

Figure 3

The self-timing page marker was designed and built in 1971 and has been used since then by Department of Education volunteers to prepare reading matter for the blind. This particular page marker was described at the Carnahan Conference on Electronic Prosthetics in 1972 by T. V. Cranmer, Director of Services for the Blind, Kentucky Department of Education, and will not be described here.

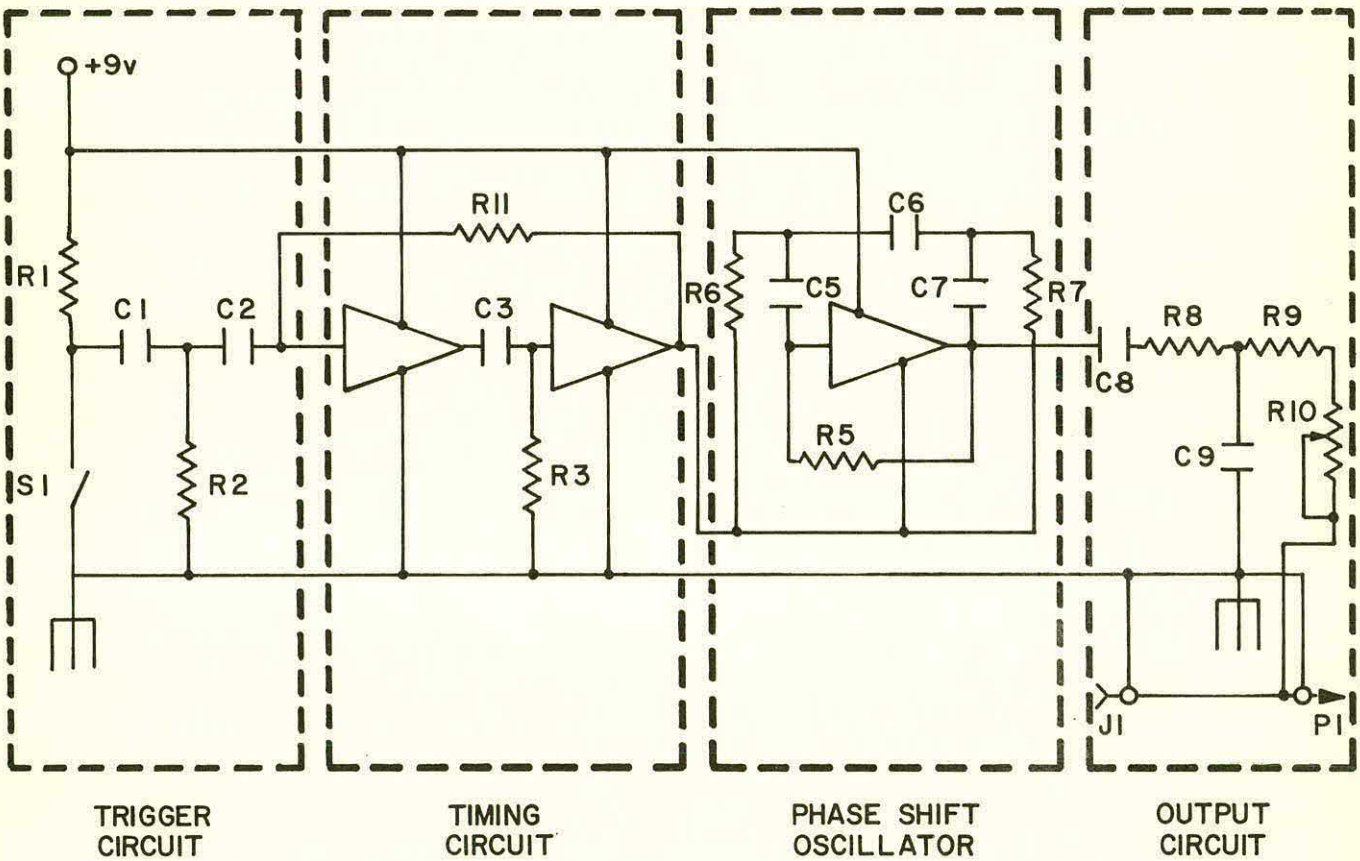
After having been used by over 100 volunteer readers for over a year, new markers were needed and it was felt that an updated design could be attained using the recently available CMOS integrated circuits. While the second generation marker uses three separate active devices, the third generation marker would use a single CMOS chip.

The integrated circuit chosen is from the digital series using complimentary MOS elements. These devices present a high input impedance, a relatively low output impedance and can operate on supply voltages from 3 Vdc to 15 Vdc. The particular device chosen is the 4007 dual complimentary pair plus inverter, shown in Figure 2.

Circuit Description

The circuit is composed of four major elements as shown in Figure 3:

1. Trigger circuit
2. Monostable multivibrator (one shot)
3. Phase shift oscillator
4. Output circuit



Schematic Diagram of Third Generation Page Marker

Figure 4

A schematic diagram is shown in Figure 4.

The trigger circuit simply supplies a negative going pulse to trigger the timing circuit. If the switch is inadvertently held down until the mark is completed, releasing the button will not re-trigger the multivibrator.

The multivibrator is composed of the first two transistor pairs in the 4007 package. The time constant is established by the .47 uf capacitor and the associated 11 megohm resistors to give a 4-5 second output pulse. Output is taken off the second transistor pair to give a ground going pulse to turn on the oscillator.

The phase shift oscillator uses the third transistor pair as the active element. It is a phase shift oscillator using a twin-T network composed of the .01 mf capacitors and the 100 K ohm resistors. Note that the high input impedance of the CMOS device allows a high impedance phase shift network so that reasonably sized capacitors can be used. Output of the oscillator is approximately 4 volts P-P with less than 1% harmonic distortion. Frequency stability, while not critical, is good. The oscillator is turned on by grounding the source of the n-channel transistor of the oscillator and by providing a ground for the phase shift network. Previous page markers controlled the oscillator by controlling the B+ supply to it. However, this created large transients requiring heavy filtering with large capacitor R-C networks. The CMOS page marker requires only moderate filtering to produce a purer output.

The output circuit is composed of a low pass filter and adjustable attenuator. The low pass filter is used to further reduce the switching transient and harmonic distortion in the oscillator. R12 is a 2 megohm ten turn trimpot to adjust the output to OVU on any of the recorders used in the Kentucky Department of Education Volunteer reading program and also to provide isolation between the output circuit and the microphone.

The page marker draws only 2 ma during a mark and less than 1 ua during idle.

Performance

In conclusion, the third generation self-timing page marker, using the latest MOS devices provides:

1. Higher reliability

    a. Lower component failure rate
    b. Greater battery life

2. Less Distortion

3. Lower switching transients

4. Lower cost

These new markers are being built for the volunteer readers in Kentucky. It is hoped that this device will encourage recording tapes for the blind to use some form of indexing or page marking.

Publications

Vol.	No.	Bul. No.	TITLES
13	2	50	Precision Voltage Ratio and Phase Shift Detector.
13	3	51	Proceedings of the Kentucky Highway Conference, February 17-18, 1959.
13	4	52	Kentucky Flexible Pavement Design Studies.
14	2	54	Hystersis Loop Analysis
14	3	55	An Investigation of the Production of Rock Dust from Kentucky Limestone for Use in Coal Mines.
14	4	56	Proceedings of the Kentucky Highway Conference, March 1-2, 1960.
15	1	57	Variation of Soil Temperature at Lexington, Kentucky from 1952-1956.
15	2	58	Mathematical Theory of a Peltier Refrigerator and a Thermoelectric Generator.
15	3	59	Thermal Analysis of the Freeze-Thaw Mechanisms in Concrete.
15	4	60	Proceedings of the Kentucky Highway Conference, March 1-2, 1961.
16	1	61	Solar Energy Measurements 1951-1960 at University of Kentucky.
16	2	62	A Survey of Components and Systems for Measuring Dynamic Loads.
16	3	63	Study of Travel Patterns in Two Lexington, Ky. Residential Areas.
16	4	64	Proceedings of the Kentucky Highway Conference, February 27-28, 1962.
17	1	65	A Short Description of Kentucky Coals.
17	2	66	Forecasting Zonal Traffic Volumes for Lexington, Kentucky, Industry.
17	3	67	Reduction of Recorder Sensitivity in Preloaded Electronic Weighing Machines.
17	4	68	Proceedings of the Kentucky Highway Conference, March 5-6, 1963.
18	1	69	Methods of Coal Storage.
18	2	70	Passive and Active Analogs with Multidisciplinary Applications.
18	3	71	A Laboratory Investigation of the Properties of Coal-Bitumen Paving Mixtures.
19	1	73	Kentucky Highway Research Program.
19	2	74	Analysis of a Proposed Solar-Earth Heat Pump.
20	1	75	Introduction to Theory of Tensors.
20	2	76	Proceedings of the 16th Annual Highway Geology Symposium, University of Kentucky, March 25-26, 1965.
20	3	77	The Photoelastic Interferometer.
21	1	78	Proceedings of the Kentucky Highway Conference, November 16-17, 1965.
21	2	79	Laboratory Investigation and Research of Three-Dimensional Flow.
21	3	80	A New Direct Method for Computing the Resistance to Flow in Open Channels.
21	4	81	Three Hydraulic Lectures Delivered at U.S. Institutions.
21	5	82	Vibration Response of a Dynamic Tension-Test Machine.
22	1	83	Strain-Rate Sensitivity Tests.
22	2	84	An Experimental Investigation of a Coupled Vortex Tube ana Radial Flow Diffuser.
22	3	85	Proceedings of the Kentucky Highway Conference, November 15-16, 1966.

Date	Bul. No.	TITLES
Aug. 1968	86	Proceedings of the 1968 Carnahan Conference on Electronic Crime Countermeasures.
Mar. 1969	87	Work Hardening Model for the Effect of Grain Size on the Flow Stress of Metals.
May 1969	88	Peak Pool Boiling Heat Flux on Horizontal Cylinders.
June 1969	89	Proceedings, 1969 Carnahan Conference on Electronic Crime Countermeasures.
Oct. 1969	90	Proceedings, Kentucky Highway Conferences.
Feb. 1970	91	Analytical Procedures for Precise Determination of Ag, Sn, Hg, Cu and Zn in Dental Amalgams.
Mar. 1970	92	Proceedings, 1970 Carnahan Conference on Electronic Crime Countermeasures.
Apr. 1970	93	A Study of the Pressure Broadening of Three Pure-Rotation Lines of HCl.
Sept. 1970	94	Proceedings, Kentucky Highway Conferences.
Apr. 1971	95	A Laser Raman Study of Two Halides of Sulfur.
Apr. 1971	96	Proceedings, 1971 Carnahan Conference on Electronic Crime Countermeasures.
Jul. 1971	97	Proceedings, Kentucky Highway Conferences.
Apr. 1972	98	Proceedings, 1972 Carnahan Conference on Electronic Crime Countermeasures.
June 1972	99	Proceedings, Kentucky Highway Conferences.
Nov. 1972	100	Viscous Hydrodynamic Instability Theory of the Peak and Minimum Pool Boiling Heat Fluxes.
Nov. 1972	101	Energy Transfer in Fur.
Apr. 1973	102	Proceedings, 1973 Carnahan Conference on Electronic Crime Countermeasures.
Jul. 1973	103	Proceedings, First International Electronic Crime Countermeasures Conference.