A Thesis Submitted For A

Master of Engineering Science Degree at

Adelaide University.

VOICE INPUT FOR THE DISABLED.

BY WILLIAM PAUL HOLMES B.E. (HONS),

MECHANICAL ENGINEERING DEPARTMENT

ADELAIDE UNIVERSITY

MARCH 1987

This thesis embodies the results of supervised

project work for the degree.

CONTENTS

	PAGE
Title Page	
Table of Contents	i
Summary	iii
Statement of Originality	v
Acknowledgements	vi
1. INTRODUCTION	1
1.1 Background	1
1.2 Outline of Thesis	6
2. VOICE INPUT - THE TECHNOLOGY AND ITS POTENTIAL FOR THE DISABLED	8
2.1 Introduction	8
2.2 Voice Input Technology	9
2.2.1 Recent and Future Trends	9
2.2.2 Principles of Operation - Isolated Word Recognition	13
2.3 Voice Input, Applications For The Disabled	25
2.3.1 Potential of Voice Input	25
2.3.2 Applications	27
2.4 Clinical Aspects - Voice Input at Regency Park	41
2.4.1 Population	41
2.4.2 Existing Methods	44
2.4.3 Potential of Voice Input at Regency Park	55
3. PROJECT METHODOLOGY	61
3.1 Introduction	61
3.2 Hypotheses To Be Tested	63
3.3 Method	66
4. SOFTWARE DEVELOPMENT	69
4.1 Introduction	69
4.2 VIM Background	70
4.3 Teacher's VIM	72
4.4 Voice Tutor	73

		PAGE
5.	TESTING	78
	5.1 Introduction	78
	5.2 Phase One - Initial Testing	79
	5.3 Phase Two - Data Input	88
	5.4 Phase Three - Speech Therapy	94
6.	DISCUSSION	105
	6.1 The Primary Hypothesis	105
	6.2 Data Input	105
	6.3 Speech Therapy	112
7.	CONCLUSION	115

APPENDIX

A1 "Implementation Of A Voice Activated Computer System For Teaching And Communication." Proceedings Of The Technical Aid For The Disabled Seminar 1985

A2 Teacher's VIM User's Manual

A3 Teacher's VIM Computer Program

A4 Voice Tutor User's Manual

A5 Voice Tutor Computer Program

A6 Experimental Data

BIBLIOGRAPHY

REAR POCKET

Teacher's VIM Disc
Voice Tutor Disc.

SUMMARY

This thesis reports on the implementation and assessment of a commercially available voice recogniser in practical situations within the disabled community served by the Regency Park Centre for the Young Disabled.

Initially a background is provided which illustrates the relevance and importance of technical aids, such as voice input equipment, for the disabled. A consideration of the development of the technology, the principles of operation and the needs of the population served by Regency Park leads to particular applications which are investigated. These applications are in the areas of data entry by voice and articulation training.

Initial testing confirmed the potential of the equipment for data entry by voice. Software was developed which made the equipment more suitable for classroom use. Trials involving disabled children were carried out in the laboratory and in the classroom. It was found that, for the trial group of children, other modalities offered greater satisfaction than voice input. Data gathered during these trials is presented as a guide to the performance of the equipment. The data is also used to illustrate a proposed methodology for the objective comparison of interfaces between the disabled and their technical aids.

The second application considered was articulation training using the voice recogniser. In this area the work centred on establishing whether the voice recogniser discriminated between sounds with sufficient accuracy to be used for articulation training. Software was developed which would be suitable for articulation training of young children. Experimental work was carried out using the software with able bodied adults saying a series of sounds simulating common pronunciation errors. The results indicated that the performance of the voice recogniser varied greatly from person to person. The data tended to suggest that the variations were made up of a relatively fixed level of substitution errors and a variable quantity of rejection errors which tend to diminish with practise. It is suggested that this characteristic may allow the equipment to be used successfully for the improvement of consistency in articulation, however further clinical testing would be required to demonstrate this.

STATEMENT OF ORIGINALITY

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university and, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except when due reference is made in the text of this thesis.

The author consents to the thesis being made available for photocopying and loan.

(W.P. HOLMES)

ACKNOWLEDGEMENTS

I wish to thank and acknowledge my supervisor, Dr. D. Radcliffe, and the head of the Rehabilitation Engineering Department at Regency Park, Dr. B. Seeger, for their advice, encouragement and criticism throughout the course. Dr. Seeger's generosity in allowing me to work within his department, interacting with his staff and taking responsibility for a project has been greatly appreciated.

The focus of the project has been the assessment and implementation of an item of new technology into a practical, clinical situation. This has been made possible by the generous cooperation of the staff at Regency Park. In particular I wish to thank Mr. P. Walker, Mr. P. Andrews, Ms. C. Olsson, Mr. P. Manson, Ms. A. Kowal and Ms. L. Stock, who have been closely involved in the testing of the equipment as acknowledged in the text.

I am indebted to the students at Regency Park who have been involved in the project for their time and patience.

I wish to thank my family and friends, particularly Ms. M. Tarrant, for the constant encouragement, support and patience throughout the project.

I wish to thank and acknowledge the Apple Education Foundation which provided funds for the computer equipment used during the project.

1. INTRODUCTION

1.1 Background

The ability to interact with and affect our environment is fundamental to the enjoyment of life. Unfortunately many people are denied this interaction by congenital or acquired disabilities. In the past there has often been a tendency to maintain the severely disabled in care institutions which have often been a further denial of interaction and stimulus. More recently there has been a growing tendency to recognise that these people have a full life to live and that society has a duty to equip them with skills and materials they may need. For example, recent legislation in the United States (Rehabilitation Act (1973) and Public Law 92-142) places a duty on the state to ensure that disabled children have equal opportunity in the school system. This duty includes the provision of services, materials and equipment which will ensure equal opportunity.

This attitude, in addition to the clear moral and ethical arguments, has a strong economic base. The minimum cost of maintaining a person in a care institution is estimated to be A$30,000 p.a. ([6] $18,250 p.a. in 1976 US$). A life-span of 50 years would place a burden of $1,500,000 on the community. In many cases, for a fraction of this cost, disabled people may gain the confidence, skills and technology to live independently, or dependently at home. There are many examples (8) of people with gross disabilities who, after receiving suitable training and equipment, have been able to support themselves and become an asset to the community's economy.

In the broader sense, motor cars and telephones may be thought of as aids for the able bodied to allow them to overcome the human body's limitations. In the same way advances in technology are being made available to the disabled to allow them to overcome their body's limitations. The abilities of the able bodied may be generally defined, and equipment designed to suit a very large proportion of the population. The operation of a motor car, for example, requires a constant stream of co-ordinated control movements from various parts of the body in response to rapidly changing stimuli which must be perceived and analysed. The fact that most able bodied adults may be trained to operate a motor car illustrates the consistency of abilities between people. This consistency has allowed a mass production and constant refinement of devices for the able bodied. The disabled, on the other hand, are often unique in their abilities. And furthermore, their set of abilities may change from time to time during their life. These characteristics make the application of technology for the disabled more difficult since the technology must be carefully selected to match the users needs and abilities.

Bleck (6) reported that the needs articulated by adult disabled persons, ranked in order of importance, are:-

1. Communication
2. Activities of Daily Living

3. Mobility
4. Ambulation.

Microcomputers are proving to be a pre-made and mass-produced function block around which specialised aids for disabled individuals can be constructed (74). Their potential is most apparent in the two highest ranking need areas listed above. Vanderheiden (75) suggests that microcomputers are revolutionising the design and development of electronic assistive devices for the disabled. However, he stresses that the major barrier for using microcomputers is the need for custom interfacing between the computer and the disabled person. He makes the point that provision of access to standard software for individuals who cannot see the video display, or cannot use the keyboard, is very difficult.

Clark's (15) opinion that "one of the primary problem areas in orthotic and prosthetic research is that of the patient device interface" is shared by many researchers and health professionals. This fundamental problem has generated an enormous amount of research into different strategies (38, 66). The diversity of solutions being considered may be illustrated by reference to examples at either end of the technological scale. Researchers have achieved promising results from strategies employing the electric signal associated with the brain's function (the electro-encephalogram, EEG) (48,71,72), and from the position and relative movement of the eye (31,63). At the other end of the scale, keyguards (68) and strategically placed switches (66) assist people with weak or spastic hand movement to manipulate keyboards.

Voice input is a strategy which has for many years been identified as an interface modality with strong potential (7,17). This modality has the additional advantage that it also has strong potential for able bodied uses. This is important because of the increased motivation to develop the technology, and ultimately the lower costs of the equipment due to economies of scale. This modality will have the potential to allow disabled people with good voice control to access computers without further customising the interface.

Speech recognition equipment has a long record of development. Experiments undertaken in 1916 (53) displayed speech as a line of varying amplitude on a strip of paper. In 1952 (20) a group of researchers developed a device which could recognise the spoken digits. In 1976 (37) a connected word recogniser was developed with a vocabulary of 1011 words which could achieve 95% semantic accuracy. However it had severe constraints on syntax and semantics and the time taken for processing was 80 times the time taken to say the statement. In 1984 IBM (33) demonstrated a large-vocabulary, real-time, isolated word recogniser. The device was able to transcribe "interoffice memo" style sentences in real time from a 5000 word vocabulary with 95% accuracy. A pause between spoken words was required, however the device was able to operate at an average discrete speech rate of 90 words per minute. 90 words per minute is approximately half the rate of normal speech. The system automatically analyses the context of the recognised string of words and corrects the spelling of words accordingly. In this way homophones like "know" and "no", are correctly entered. The IBM research group has recently reduced the size of hardware required so that it resides in the back of the AT range of personal computers. The group predicts that a personalised vocabulary of 15000 words would require the user to spell 1% of words, and that an increase in computing speed of a factor of 6 to 10 will allow continuous word recognition in real time. There are some reported cases (8) of highly successful applications of speech recognition by the disabled where the person's other options were severely limited. As the technology improves, the number of people who may benefit will increase because the potential of the technology will exceed their other alternatives. Reference to the historical development of voice recognition equipment shows that progress has been achieved in relatively small increments, and generally has been aimed at the able bodied user. Therefore there will be no clear date when more members of the disabled community will benefit from the technology. Health professionals must therefore keep abreast of advances in the technology and experiment with available equipment to assess potential.

This thesis reports on a project which takes up the challenge of assessing the potential of a speech recogniser in a disabled community. The underlying philosophy has been to take the equipment out of the laboratory and assess it in the real world. The thesis reports on work involved in customising and introducing the equipment and quantifies the results. It is expected that the thesis will assist health professionals wishing to gain an appreciation of the strengths and weaknesses of available equipment. The thesis also provides information for researchers in voice recognition wishing to gain an insight into the special problems of the disabled.

1.2 Outline of Thesis

Chapter two develops a background against which the relevance of the succeeding chapters may be seen. The classification and principles of operation of voice recognition equipment are discussed. The existing applications of voice input equipment are discussed and applications for the disabled are considered within five categories. A proposed methodology for comparing data input modalities is also presented. The needs and existing alternatives for the students at Regency Park are then considered. This allows the selection of the two applications of voice input offering greatest promise.

Chapter three outlines the methodology adopted for the assessment and implementation of the voice input equipment. A primary hypothesis and six secondary hypotheses are developed. These hypotheses form the basis of the experimental work carried out during the thesis. The chapter also discusses the three phases of the project.

Computer software was modified and developed during the project. Chapter four describes these developments.

The experimental work is described in detail in chapter five. The chapter is divided into the three phases of the project and discusses the experiments under the following headings:

(a) Hypothesis tested

(b) Subjects

(c) Design

(d) Apparatus

(e) Procedure

(f) Results

(g) Interpretation.

Chapter six discusses the primary hypothesis and takes an overview of the interpretations of results in chapter five.

The conclusions drawn from the project are presented in chapter seven.

2. VOICE INPUT - THE TECHNOLOGY AND ITS POTENTIAL FOR THE DISABLED

2.1 Introduction

This chapter develops a background against which the relevance of the succeeding chapters may be seen. Initially, Section 2.2 discusses the classification of voice recognition equipment giving an indication of the wide variety of sophistication and performance. The principles of operation of voice recognisers are then discussed and a historical perspective is provided to aid the reader in understanding the developing trends in the technology. Particular emphasis is placed on describing the type of recogniser used during the project work associated with this thesis. Some of the principles behind the recent developments in the technology are discussed to provide an insight into future trends.

The potential of voice recognition technology for the disabled is considered in Section 2.3, initially in general terms. Applications of voice recognition for the disabled are divided into five categories and documented applications are considered.

Section 2.4 describes the function performed by Regency Park and the methods adopted. The population served by Regency Park is described in general terms and the man machine interface is discussed. Existing methods used by the severely disabled to control devices provided for their service are discussed. Finally the relevance and potential of voice recognition equipment at Regency Park is discussed under the five categories developed in

Section 2.3. Directions for development of voice recognition equipment at Regency Park are therefore defined.

2.2 Voice Input Technology

2.2.1 Recent and Future Trends

The advent of the large scale integrated circuit has revolutionised speech recognition technology so that today there is a very broad range of equipment available. The various types may be generally categorised by a brief reference to their sophistication. For example **speaker dependent, isolated word recognisers** recognise, or are activated by words separated by a distinct pause, spoken by a person who has previously trained the device with his own voice.

The simplest type of recogniser is the **single sound recogniser.** The keyring that beeps when the user whistles is an example. Even this simple technology may be applied in a useful fashion, for example a computer program which displays a moving train when the continuous "Shhh.." sound is made. If the sound is interrupted then the train stops. It is suggested that this may aid Speech Therapists in motivating some clients. It is conceivable, also, that this type of single sound device might be used by a disabled person to, for example, switch the lights on or off.

The **speaker dependent isolated word recogniser** represents the next class of recognition devices. Again there is a wide range of equipment available representing this class. Devices are available (32) for approximately $200 which activate one of ten switches depending on which of the ten previously trained words are detected. Several recognisers (36,46) have become available in the market place during the last ten years which cost between $500 and $2500 and function as "keyboard emulators". These devices are generally combined with a personal computer and generate signals in response to spoken sounds. The signals are interpreted by the computer as a series of keystrokes so that the recogniser may be used as an alternate keyboard. These devices generally recognise the sound based only on the signal received for the particular sound. Their accuracy, therefore, tends to fall off rapidly for large vocabularies, or for vocabularies containing similar sounds. They therefore tend to be restricted to less than 100 words, although some examples have inbuilt logic and memory which allows the user to rapidly switch between several vocabularies allowing ready access to a vocabulary of 1000 words. The speaker dependent isolated word recogniser is the most sophisticated variety of easily available, inexpensive, recognition equipment. Consequently it is the variety which has been most widely applied in the community. Applications for the able and disabled will be discussed in Section 2.3.

The last paragraph noted that the isolated word recogniser is generally limited in vocabulary size due to the inaccuracy inherent in techniques which consider the sound signal only. Attempts are being made to overcome this problem by considering the contextual and syntactic information in the string of words. Devices currently being developed (29,30,33) display a sophistication far in excess of devices already described. These devices have already successfully demonstrated recognition of words from a 5000 word vocabulary, at 95% accuracy, at rates of entry of 90 words per minute. This "state of the art" equipment is housed in the IBM AT personal computer. This device requires the user to initially train his voice onto the machine and to provide a distinct pause between words.

Jelinek predicts (33) that an increase in computing speed by a factor of 6 to 10 will allow **speaker dependent continuous word recognition.** This will allow a speaker who has previously trained the device to speak at a normal rate without pausing between words. They suggest that a personalised vocabulary of 15000 words will result in the user having access to 99% of words required for normal business memoranda. This will lead to a situation where a relatively untrained person could prepare typed material at a rate of 150 words per minute. It may be seen that the potential market for such a device is enormous. The size of the potential market will stimulate more intensive research and ultimately the economies of scale will reduce the cost of the equipment.

The recognisers described above have generally been speaker dependent. Researchers are also very active in developing **speaker independent isolated and continuous word recognisers.** Generally the problem of speaker independence is more difficult since the recogniser must allow for variations associated with sex, age, dialect and other individual characteristics. Despite these problems, progress is being made. The French recently announced plans to trial telephones with voice activated dialling, which is an example of a speaker independent, isolated word recogniser. Myers (47) demonstrated a connected word speaker independent recogniser in 1982. The device achieved a 98% correct word recognition rate for four people who spoke to the recogniser over a telephone. No previous training was required, however the vocabulary was short and the syntax highly constrained.

As well as the categories already mentioned, recognition equipment may be further divided by the following self explanatory qualifications (18):-

(a) Small vocabulary and constrained syntax/large vocabulary and natural language syntax.
(b) low/high information bandwidth
(c) trained/untrained speaker
(d) quiet/noisy environment
(e) co-operative/uncooperative speaker.

The goal of high performance speech recognition approximating the rates and accuracy naturally achieved by human listeners is still a long way off. However, substantial progress has been made resulting in commercial systems which point the way to the future.

2.2.2 Principles of Operation - Isolated Word Recognition

Many schemes exist for the automatic recognition of sounds. The following discusses a common approach for sounds which are analysed without reference to adjoining sounds.

Word, syllable or phoneme recognition is performed by comparing parameters derived from incoming speech (feature extraction) with stored descriptions of candidate recognition items (template matching) as depicted below (17).

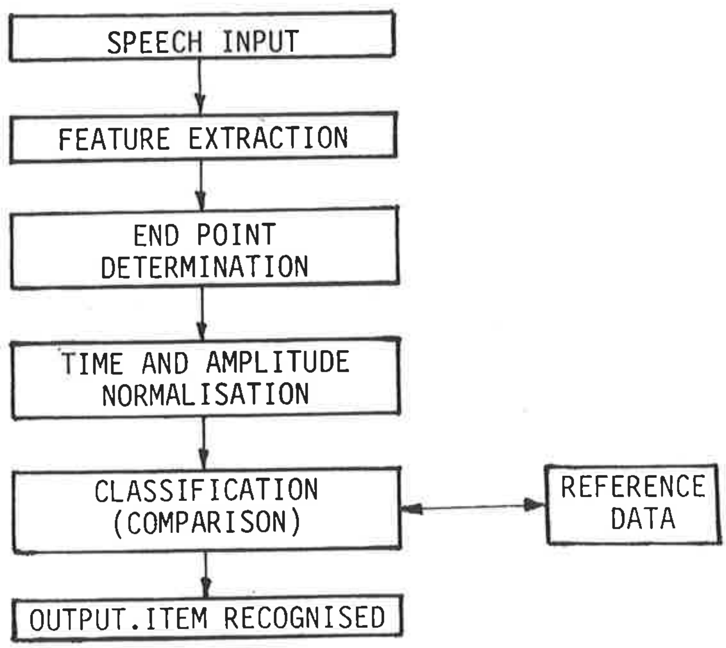


FIGURE 1. THE STEPS INVOLVED IN SPEECH RECOGNITION

Feature Extraction

The incoming speech signal contains an enormous amount of information. The "feature extractor" attempts to isolate distinctive characteristic in the sound which will allow the computer to identify the sound. Signal processing techniques and a basic knowledge of human speech processes are used by the recogniser designer to select the characteristics which will be extracted. Designers have adopted many different characteristics features of the voice signal. The Bell group which developed the spoken digit recogniser in 1952 (20) monitored the zero crossing frequency in the speech signal above and below 900Hz. This gave an approximation of the time varying first and second formants. When the researchers displayed the time plot of the second formant against the first formant they found a unique pattern for each spoken digit. They found there was sufficient repeatability to allow identification by visual inspection, and then developed electrical circuits to recognise the patterns automatically.

A number of modern recognisers record the energy of the signal of several fixed frequencies during the utterance (53). It may be seen that this is similar to the Bell group's (above) system since the technique will effectively detect the formants but goes further since it monitors the strength of the formants and gathers information about other frequencies as well. This is particularly important in recognising unvoiced fricative sounds which do not have distinct formants. In these sounds, produced by a constriction in the vocal tract, the spectrum is characteristically continuous with a wide range of frequencies (45).

Information on the formant frequencies is by no means the only information which may be gained in this way. Zwicker (82) discusses the extraction, from a set of filters, of the time varying parameters which he labels:- "loudness, pitch, roughness, timbre, and subjective duration". In another study Reddy (57) lists 21 parameters.

Atal noted (3) that the transformation of the acoustic data into spectral form (described above) has serious limitations. In particular the traditional Fourier analysis methods require a relatively long speech segment to provide adequate spectral resolution. Consequently, rapidly changing speech events cannot be accurately followed. The techniques also generally provide little information about the spectrum between pitch harmonics due to the quasiperiodic nature of voiced speech.

Atal considered that the first step in signal analysis should be the construction of a model of the signal. The model he proposed comprised two parts; a sound generator followed by a sound shaper, or filter. The generator may take the form of a low frequency periodic signal (representing the action of the vocal cords in voiced speech), or of a relatively random mixture of high frequencies (representing the unvoiced sounds produced by constrictions, or sudden releases of air in the vocal tract). The sound shaping filter is a model of the shape of the vocal tract which attenuates various frequencies. Recognisers using this model tend to analyse the speech signal in two stages. In the first stage the recogniser determines the nature of the excitation frequency (i.e. voiced or unvoiced). In the second stage the recogniser establishes a number of parameters which define the filter at the particular time.

The recogniser described by Fright et al (25) uses the method described by Atal. The first stage is accomplished by extracting three measures of the speech energy:- (i) the short time average, (ii) the low frequency short time average (200-1000Hz) and (iii) the high frequency short time average. The speech is termed unvoiced if the high frequency short time average is significantly greater than the low frequency short time average, otherwise the speech is termed voiced. The second stage of analysis uses linear predictive coding (LPC) to determine 10 parameters every 20ms throughout the voiced or unvoiced segment. These 10 parameters define the vocal tract shape (filter model) at each time interval. LPC is based on the premise that samples of voiced speech are linearly predictable in terms of past speech samples. A least squared, or similar method, is then used to define parameters which minimise the error in the last sample when compared with a predicted value based on previous samples.

The strategies above utilise models of speech production. Zue (81) points out that it may be more appropriate to consider and model the auditory system since our problem relates directly to the recognition of sound. He reports on work with filters which emulate the "frequency analysis" carried out by the ear (4) which tends to give a better frequency resolution at the low frequency end of the scale. In addition, he goes on to report that the action of nerve fibres in the ear tend to enhance peaks at the format frequencies. A detection algorithm which models the synchronous firing in the auditory nerve illustrates the way in which the formant peaks, and concentration of spectral energy are enhanced. He suggests that a recogniser which is "tuned" to receive speech, and enhances speech-like sounds, will tend to achieve better performance in noisy environments.

End Point Determination

The recogniser must be able to correctly detect the ends of the subject sound. Failure to do this will lead to the situation where the reference sounds are compared with an incomplete sound or a combination of two or more sounds.

Isolated word recognisers require that only a single word is spoken at a time. Not only does this assist in end point determination, but also decreases the variability in pronunciation due to preceding and succeeding words. In some recognisers this is also necessary since the equipment can either extract features or normalise and classify, so that a pause is necessary for the processing of information. End point detection is further complicated by the presence of "stops" in sounds, and the presence of breathing, lip smacks, clicks and other spurious sounds after the sound is completed. The recogniser designer overcomes these obstacles, to some extent, by specifying acceptable lengths of sounds and the length of pauses between sounds. For example the recogniser used in the study associated with this thesis, the VIM (76), will only accept sounds which continue for at least 0.1 seconds but no longer than 1.25 seconds and with a pause between sounds exceeding 160ms.

End point detection is a major problem for connected word recognisers. Speakers tend to run sounds together (coarticulation) making it difficult to recognise individual sounds. This may lead to the situation where a "stop" within a sound exceeds the gap between words. A number of connected word recognisers attempt to isolate phonemes (57) or syllables. End point detection of these segments becomes more complicated and depends on changes in the nature of the signal (4,22,27,41,59). The difficulties inherent in end point detection have led some researchers to develop strategies which consider many possible end points so that the real end point will not be missed. For example Sakoe (62) developed an algorithm which matched reference words with possible subsets of the input signal to produce many "possible" words. These "possible" words were then tested at the phrase level for consistency.

Time and Amplitude Normalisation

This section has been describing recognisers which depend on the comparison of the incoming speech signal with a stored template. Clearly, variations in the speed of production of the sound, and the loudness of the sound will lead to a failure to correctly compare an incoming sound with the stored template unless some correction is applied. The simplest form of correction, or "normalisation", amplifies the signal so that its average strength is a predetermined value, and divides the signal into a predetermined number of divisions. The values of the characteristic features are then recorded at the end of each time interval. The number of time intervals corresponds with the number of divisions in the stored template so that there is a one to one correspondence at the comparison stage.

The technique described in the last paragraph is limited because the relative timing and strength of characteristic features within the sound alter when the speaker alters the speed or amplitude of pronunciation. The technique assumes that a feature which appears at, say, the midpoint of the utterance will always occur at the midpoint, irrespective of the way in which the word is pronounced.

The technique of dynamic time warping (DTW) is used extensively in isolated word recognisers to overcome the limitations described above. In general terms the time scale of the incoming signal is altered so that the signal is made to fit the templates. The combination of best fit is then taken to be the correct combination. Effectively the normalisation and classification activities are combined. DTW will be discussed further below when classification is considered.

Classification (Comparison)

Classification refers to the stage where the characteristic features of the incoming signal are compared with the characteristic features of the stored template. A decision is made as to whether the incoming signal is sufficiently similar to one of the reference patterns, if so then the recogniser assumes the incoming signal is the same as the reference patterns and generates the appropriate output.

This comparison may be achieved in many ways. The recogniser developed by the Bell group in 1952 (20) relied on a relay, capacitor network to increase charge on capacitors when close fit between the incoming signal and stored template was achieved. The capacitor with the most charge at the end of the analysis was therefore associated with the best match between sound and template. This capacitor was then identified by means of a relay, gas tetrode network.

Analog circuits are being developed (45) however most recently developed recognisers are digital. Again there has been a wide range of techniques employed. Clark (15) described a very simple technique which is used in a wheelchair controller. He divides each incoming sound into eight segments, each of which is characterised by 8 bits of information. Therefore each sound is characterised by an 8x8 matrix of digital information (64 bits total). He then matches the incoming 8x8 matrix with each of the 8x8 template matrices. If a single template is found that matches 58 or more bits of the incoming matrix then the wheelchair command associated with the matrix is activated. If more than one template is found to match 58 bits then the command associated with the highest number of matched bits is activated. Clark reports successful use of the recogniser although it is limited to recognise only 8 distinctly different words. The same basic technique is used in many recognisers, including the one used in the experimental work associated with this thesis (76). However, improved performance is achieved by analysing a greater number of segments and by recording the value of the amplitude of 16 frequency bands at the end of each segment.

The methods described in the last paragraph do not account for the problem, identified earlier, of the relative displacement, with respect to time, of features in the sound. The speed of modern digital devices allows the use of algorithms which perform a search of the various ways that the time scale of either the incoming signal, or the reference template, may be distorted to achieve a best fit. The many variations of this theme are described and compared in several references (21,12,34,35) however the figure prepared by Levinson (39), and reproduced below (Figure 2), gives a clear idea of the concept.

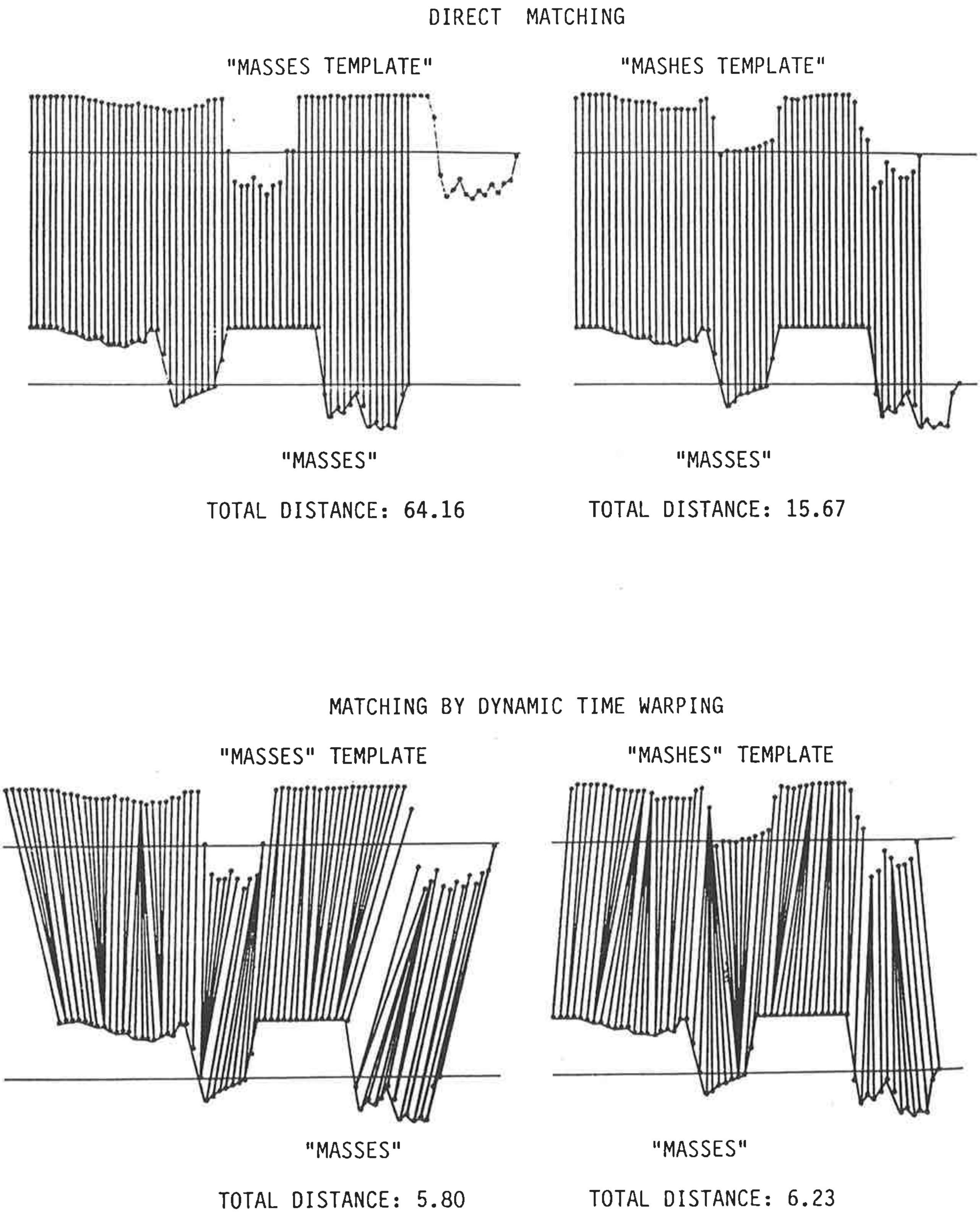


FIGURE 2. This figure compares the direct matching of characteristics of sound signals with the dynamic time warping method.

Figure 2 uses the distance between the plot of the characteristics of the incoming and template sound as a measure of overall fit. The non-uniform time alignment used in the dynamic time warping method allows for variations in the rate of speech and in the relative lengths of the vowels and consonants in a word.

The Recogniser as a Whole

Most isolated word recognisers perform the four tasks in Figure 1 to achieve recognition of an incoming sound signal. The four tasks (i.e., 1. Feature Extraction, 2. End Point Determination, 3. Time and Amplitude Normalisation, and 4. Classification) have been discussed in general terms to explain the tasks more fully, and to illustrate the various methods used to carry out the tasks. The following brief description illustrates how the tasks described work together to perform the analysis and recognition of speech signals.

The speaker's vibrating vocal cords cause pressure fluctuations in the air. A microphone converts these pressure fluctuations to an electrical signal which is amplified and monitored by a speech recogniser. The first function of the recogniser is the extraction of relevant information from the signal. Hendtlass (28) suggests that there are 48000 bits of information per second of typical speech. The speech recogniser attempts to discard information not essential for word recognition by concentrating on particular time varying parameters in the signal, such as the formant frequencies. By doing this the recogniser eliminates less important information and reduces the amount of data which must be handled. Clark (15), for example, reduces sounds (words) to a total of 64 bits of information. While extracting information from the signal, the recogniser constantly monitors the parameters for indications of the endpoint of the sound. This is particularly important because failure to accurately detect the endpoint will lead to the situation where the recogniser will be comparing a pair of words, or a partial word with its library of single words. Clearly the recogniser will be unlikely to achieve the correct comparison in these circumstances unless special techniques are used. After detecting the endpoint, the recogniser has a string of data representing the time varying parameters of the incoming sound. To accurately produce the word associated with the given string of data the recogniser must first account for the speed and amplitude of pronunciation. After this "time and amplitude normalisation", the recogniser compares the string of data describing the incoming signal with the recogniser's library of stored data. If the incoming data string matches a particular data string in the library then the recogniser retrieves the word associated with the library string and the recognition is complete. The recogniser may then activate a switch, output the word, or a character, or perform some other function depending on the design of the recogniser and the word recognised.

2.3 Voice Input, Applications For The Disabled

2.3.1 Potential Of Voice Input

Many researchers have identified the potential strengths of the voice in the man-machine interface (7,17). The list of apparent advantages includes:-

(a) Voice is a natural means of communication.

(b) Voice input is good for "hands busy", "eyes busy" situations. In certain cases of disability, it is the only channel available.

(c) It requires no physical linkage (in contrast to switches, joysticks, pointers or EMG skin electrodes).

(d) A wide range of control freedom.

(e) Optimisation for the personal needs may be achieved by software modifications.

(f) A potential for high speed communication.

(g) Ease of access for telephone or radio.

(h) Advantages in situations of poor lighting or visibility or for users with poor eyesight.

(i) Psychological advantages - fun, motivating, user has a sense of control by whim.

However, these potential advantages are often not achieved in available, inexpensive recognisers. The following limitations were identified by Boonzaier (7):-

(a) The inability to recognise the same words repeated with different pitch, inflection, speed, loudness, phonation and connectedness, may impose too great a demand for consistency on the user.

(b) Recognisers are generally speaker-specific. They must be taught the user's pronunciation, and the user must pronounce the words in the same way when using the recogniser.

(c) Increasing the size of the vocabulary of the recogniser increases the possibility of the recogniser confusing two sounds and also increases the time taken by the recogniser to compare sounds. The size of the vocabulary is therefore restricted - in most inexpensive recognisers the active vocabulary is less than 100 sounds.

(d) Recognisers generally fail to use the syntactic information implicit in formed speech. Human listeners tend to anticipate following words, and review past words with reference to learned grammatical rules. Grammar dramatically decreases the number of appropriate words which might be uttered. The human listener subconsciously discounts improbable sounds and replaces them with more probable substitutes.

(e) Similarly, the human listener uses the semantic and contextual framework of a string of words to discount improbable sounds and replace them with probable substitutes. The sense of the following words assists in the selection and confirmation of the substitute. Inexpensive recognisers take words uttered in isolation and do not consider the adjoining words.

(f) Visual cues perceived by the human listener provide information on the "place" in the vocal tract where the sound was generated and the "manner" of production. The "body language" displayed by the speaker may also be used by the listener to set the contextual framework of the utterance. This visual information, which is not available to voice recognition equipment, assists the human listener to assess the perceived sound.

The limitations discussed above are common to several speaker dependent, isolated word recognisers produced by manufacturers as interfaces for personal computers since the mid 1970's. One of these recognisers is the subject of the experimental program associated with this thesis. More recently (33) researchers have developed strategies and equipment which radically reduce the effects of the limitations (a) to (e) listed above.

In general terms, the exploitation of voice recognition follows a common theme. The recogniser (or interface) identifies the utterance (the user's input) and passes a signal to another device (generally a computer) which controls the required output. The following subsection illustrates the ways in which this theme may be exploited.

2.3.2 Applications

This subsection reviews published applications of voice input. The review is not intended to be exhaustive but is intended to illustrate the scope of applications being considered. Applications are divided into five areas, i.e. Data Entry, Environmental Control, Speech Therapy, Aids for the Deaf, and for Translation.

Data Entry

The entry of data by voice rather than by hand is perhaps the most obvious use of voice input. The following five authors refer to applications where voice input has been competitive with hand input.

(a) Martin's (42) 1976 paper titled "Practical Applications Of Voice Input To Machines" gives several examples, which he divides into four groups:-

(i) Quality control and inspection:

- Television and faceplate inspection.
- Inspection of pull ring can lids.
- Automobile assembly line inspection.
- Receiving inspection. The recording of information on goods entering a factory.

(ii) Automated material handling:

- Voice controlled package routing system.

(iii) Parts programming for N.C. machine tools:

- Allows factory personnel unfamiliar with programming techniques to program a N.C. tool with spoken commands.

(iv) Government applications:

- Controls for high performance aircraft cockpit simulators.

- Air traffic control.
- Entry of cartographic and bathometric data.
- Aids for the handicapped.

Martin states that in several of the above cases there has been increased operator productivity and reduced error rates compared with keyboard entry of data. He also reports that average data entry rates of 30 to 70 entries per minute have been regularly achieved throughout 8 hour working days, and peaks of 120 entries per minute have been recorded.

(b) Rollins (61) reported in detail on two industrial applications of speech input of data where the operator's hands were generally occupied with other tasks. The applications were:-

(i) A baggage handling facility where voice was used to define the destination of an article.

(ii) An inventory receiving facility where the incoming article's weight, identity numbers and description were entered by voice.

The investigators found that in general terms voice entry was comparable to keyboard entry, however operator attitude and ability often made speech input either significantly more or less efficient than keyboard entry. They also noted that some potential advantages of voice entry were lost since the systems were set up for keyboard entry.

(c) Birch (5) reported on work by the Royal Aircraft Establishment which is developing a "cockpit control system" activated by voice which manipulates radio and navigational equipment. The equipment is currently being used in a BAC-111 aircraft on a trial basis and is expected to be used in military aircraft in the near future.

(d) Spine (70) refers to several authors when he listed the following applications of voice entry of data:-

(i) Distributed computer networks command and control.

(ii) Air traffic control.

(iii) P-3C aircraft command and control.

(iv) A warfare simulation trainer.

(v) Printed circuit board design.

In the situations described above, voice is competitive with keyboard entry because the user's hands are otherwise occupied. Similarly voice input for data entry has applications for disabled people with good voice but poor hand control. This is illustrated by the following:-

(a) Richards (58) reports that people with impaired hand function are able to use voice input to complete a psychological test, the Minnesota Multiphasic Personality Inventory. The method is substantially cheaper than employing a scribe for the 2 hours required to complete the test. Richards claims a 97% accuracy and states that the 3% errors randomly distributed through the test do not affect clinical profiles significantly.

(b) Bowe (8) reports that a C1 quadriplegic has been able to use voice input of data to operate a home based computer. He has been able to obtain employment as a computer assistant and worked five hours a day for the National Institute of Health.

(c) Fu (26) reports that Boeing has devised a speech controlled workstation for a quadriplegic person who is able to work in a "business programming environment". The group provided an extended vocabulary set which enables the user to employ graphics, support multiple programming languages, access different networks and control a personal telephone system. The system also allows the operator to manipulate the immediate environment which is described in the environmental control subsection below.

(d) Rizer (60) refers in general terms to a quadriplegic student undertaking a computer programming course at the Maryland Rehabilitation Centre using voice input for computer access.

(e) Creasey (16) reports on the development of a voice controlled, data entry arrangement for tetraplegics. The system used a two and three dimensional selection procedure (a form of encoding - Section 2.4.2) to obtain reliability with a small vocabulary. This reduced the data entry rate to 7 characters per minute.

(f) Fried-Oken (23) reported that a 10 year old C1-C2 quadriplegic displayed increased motivation to learn following use of voice input. This was despite poor accuracy due to fluctuations in vocal productions, interfering noise from the uncuffed tracheostomy, and unreliable recognition of a high pitched voice.

Environmental Control

The subject of environmental control by voice overlaps with the data entry application discussed above because data must be entered to produce the output which controls the environment. The areas are discussed separately here because the two subjects tend to have different basic requirements. Data entry requires quick operation and a large vocabulary, but will generally tolerate some errors. Environmental control units on the other hand require a high degree of accuracy but generally require only a small vocabulary and tolerate slow operation.

Environmental control with speech input has been a popular area for research because current inexpensive recognisers are able to meet the general requirements. The required accuracy may be achieved by "syntaxing", or "error trapping" where two or three sounds are entered for one action. The second and third sounds allow the user to confirm or cancel the initial sound. Some areas of interest, particularly wheelchair control, have a low tolerance for errors and require fast operation. The relatively small required vocabulary assists the many researches exploring wheelchair control, however reference to the literature did not reveal any examples in regular or normal use.

The following illustrates applications of voice in environmental control:-

(a) The speech controlled workstation (26) described in the data entry discussion above also provides an ambitious control of the local environment for the user. The developers intended that the quadriplegic operator of the workstation would not require any human assistance in normal tasks. A voice controlled robotic system has therefore been provided which performs the following tasks:-

(i) Retrieves and displays books, turns pages and returns the book to the storage location.

(ii) Opens and closes filing cabinets, inserts and retrieves documents and assists the operator in searching the file.

(iii) Retrieves, loads, unloads and stores floppy discs.

(iv) Loads and unloads the printer and displays printed sheets.

(v) Performs miscellaneous tasks such as transferring material to the waste paper bin, to the user's lap, the user's back pack or the in/out tray.

Fu reports that the system was operational in April 1986 and that the quadriplegic "operator is completely independent of supportive aid from co-workers".

(b) Brown (11) presents a case study of a 42 year old, mentally retarded (IQ19) woman who possessed no self help skills and almost completely unintelligible vocalization. A voice activated device was provided to allow the subject to operate a videotape of classical stories, a vibration message pad, a radio, and a videotape of the subject's family. "The subject learned the relationship between her vocalizations and the activation of environmental appliances and exercised demonstrable control over her environment for the first time in her life." Brown reported several positive gains in the personal development of the subject during the study.

(c) Damper (18,19) discusses the use of voice input, particularly for environmental control and describes in detail a dialogue design whose syntax ensures that errors due to inaccurate sound recognition are suppressed.

(d) Waby (78) describes in general terms the Voice Connection (77) environmental control unit and its impact on a 23 year old quadriplegic person at Melbourne's Austin Hospital.

(e) Many researchers have reported on the development of voice controlled wheelchairs. Soede (69) discusses experiments designed to determine a command set for wheelchair control. Miller (45) discusses the performance of an inexpensive analog sound recogniser and Clark (15) describes a simple digital recogniser and associated trials of wheelchairs driven through a maze using voice. Clark's wheelchair control also employs an emergency stop switch activated by head movement.

Speech Therapy

Speech recognisers depend on extracting parameters from the spoken sound which characterise the sound. It would seem that such a device could be used to:-

(1) Check that a sound has been spoken correctly by ensuring the parameters are approximately the same as the parameters of a target sound.

(2) Show which parameters deviated from acceptable limits during a mispronunciation.

It is suggested that speech therapists could apply the equipment in the following two areas:-

(1) Articulation may, in some cases, be improved by repetition of the correct pronunciation of a sound which the client finds difficult. Articulation drills use this

technique. Generally, however, a therapist is required to constantly monitor, guide and encourage the client through the drill. Alternatively, when the client reaches an acceptable stage a recogniser might be used to monitor the sounds produced and motivate the client. This would free the therapist to attend to other clients and would allow the client to perform articulation drills between visits to the therapist.

(2) A speech recogniser detects a mispronunciation when the detected parameters defining the sound deviate from a "target" sound by an unacceptable amount. It is conceivable that the parameter deviation could be analysed and that a computer could show the speaker what caused the mispronunciation. Computer graphics might be used to illustrate the error. For example, diagrams of the vocal tract for the target sound and the produced sound could be produced to demonstrate an error due to place of sound production.

Research into the areas discussed above has not been extensive, however some work has been done. Fried-Oken (24) reported that a mildly dysarthric speaker improved articulation due to articulation drills monitored by a speech recogniser. The subject used four words to play a computer game which tended to increase motivation and attention span. A 125% increase in attention span was noted. In another study, Schmitt (64) noted improved articulation in a severely disabled subject who has the ability to produce only semi-intelligible utterances. In this case the improved articulation appeared to be due to improved breathing co-ordination which was encouraged by the repeatability of the speech recogniser.

Fright (25) has developed a speech recogniser which may be connected to an IBM PC computer which displays graphs of the characteristic parameters of both the "target" and incoming sound. He reports that they also plan to display speech energy, pitch and vocal tract shape. These graphics might well be used by a Speech Therapist to show the client more clearly the cause of their error in pronunciation.

Aids For The Deaf

Speech recognition is a major concern for the deaf. In the far distant future speech recognition equipment may develop to the point where conversational speech may be analysed and displayed on a deaf person's glasses as text. In the foreseeable future however technology will attempt to provide clues to assist the deaf in lip reading. Technology for the deaf often duplicates techniques in other areas of speech recognition research, particularly in feature extraction. Equipment for the deaf seeks to extract features from the incoming speech signal and present information as either visual or tactile stimulation. Lip reading tends to provide information on place of sound production, additional information on voicing and manner of production has been shown to assist the lip reader. The following summaries illustrate the variety of approaches being taken.

(a) Brooks' (9,10) experiments with an array of vibrators attached to the subject's arm revealed that the subjects were able to learn to recognise 250 words by reference to tactile information only. The equipment employed 16 filters with centre frequencies between 200 and 8000Hz. The amplitude of 16 vibrators was then modulated in proportion to the energy detected by the corresponding filter channel. This approach has many similarities to the feature extraction technique employed by many single word recognises including the recogniser used at Regency Park as discussed later in this thesis.

(b) Traunmuller (73) devised a series of ten light emitting diodes (LED) which show the ratio of energy in the high frequency (2.2kHz) versus the low frequency (680Hz) part of the speech signal. The ratio varies with time as the word is spoken leading to characteristic changes in the LED array. Characteristic and consistent changes have been associated with various stops, voiced continuants, unvoiced stops and unvoiced fricatives.

(c) Pickett (52) reports on experiments with the "Upton eyeglass speechreader" which tends to indicate improved lip reading performance. The eyeglasses display a pattern of five lights to the wearer which indicate when any of the following four conditions are encountered:-

(i) unvoiced high friction for greater than 80ms.

(ii) low frequency friction (2-3 kHz)

(iii) brief burst of less than 80ms after silence i.e. plosives

(iv) low frequency murmur.

(d) Miller (44) reports a 0 to 30% improvement in lip reading performance when vibration applied to the fingers conveys additional information on the following phonetic features:- voiced/voiceless, continuant/interrupted, and nasal/oral.

(e) Martony (43) reports some improvement in lip reading performance when information on voicing and "stops" is provided either visually or as vibration on the skin.

Translation

Unintelligible speech is clearly an important barrier to communication for many people. This may occur even though the person produces intelligent and consistent utterances. A relatively small amount of research is being done to try to alleviate the problem by creating devices which analyse the person's speech and then produce synthesized speech recognisable by the wider community.

Translation of speech by machine is not a new idea. Olson (50) demonstrated a device in 1962 which recognised a limited number of English words and then pronounced them in French, German or Spanish.

Rahimi (54) identified two approaches to this translation problem. The first is to have a computer recognise the dysarthric speech on a word- by-word basis and then to synthesize a more generally intelligible equivalent. The second approach is to have a computer systematically modify the speech signal without an attempt at actually recognising the speech.

The first of these approaches has been used by Schmitt (64) who provided a severely disabled, visually impaired individual with a computer system which can both recognise and synthesize speech. The subject is able to consistently produce semi intelligent utterances which are used to prepare sentences. These strings of words may then be sent to the synthesizer which produces speech intelligible to the wider community. This procedure is slow and laborious, however it should be remembered that the subject has no other way of preparing and storing messages for spoken communication with strangers.

Fright (25) reports that his group has produced an analysis/translation/synthesis scheme which has produced encouraging results. A portable device is being constructed which will run in real time, overcoming the disadvantages inherent in Schmitt's (above) approach.

2.4 Clinical Aspects - Voice Input At Regency Park

2.4.1 Population

Regency Park is a school for disabled children. The school generally caters for children from kindergarten through to late primary school, and to a lesser extent secondary school. The school offers curricula which are similar to other South Australian schools and attempts to equip the students with skills and abilities which will enable them to leave Regency Park and integrate into mainstream schooling. The goals of mainstreaming (13) are:-

(a) to permit the child to attend a local school, thus reducing transport problems,

(b) to maximise educational expectations and performance, and

(c) to allow social interaction with able-bodied peers.

Children who cannot currently attend a regular school due to their disability, but demonstrate that they may, with assistance, be able to attend a school in the future, are eligible candidates for Regency Park. In some cases this integration process is not possible and, if further tuition is impractical, these children may go on to care type institutions. A great deal of compassion is obviously required in this situation, particularly when degenerative disabilities are concerned.

The fundamental criterion for entrance to the school is an expectation that the child will benefit from the school. The actual type of disability is of secondary consideration, so that the range of diagnosis is very broad. Furthermore, the severity of disability within a single diagnosis may vary greatly. For example, cerebral palsy, which is the most common diagnosis, may be exhibited in some cases as spasms of some or all muscle groups such that conscious, predetermined movements are impossible (cerebral palsy, spastic quadriplegia). In other cases cerebral palsy may be characterised by a lack of spatial orientation with no tendency to spasm (cerebral palsy with marked ataxia). The range of diagnosis is also very broad often referring to a congenital weakness of particular muscle groups (for example:- Arthrogryposis) or of skeletal mal- development (for example:- Osteogenesis imperfecta) which prevents the student from safely moving about or manipulating objects in the normal way due to risk of, or treatment of bone damage.

The time of onset of the disability is also variable. The disabilities discussed above are all apparent from birth. Other students, however, have lived a normal life until a road, or other accident leads to, for example, a head injury or spinal damage. Injuries to the central nervous system may cause changes of varying severity to many abilities including speech, co-ordinated movement, cognitive ability, memory, interpersonal skills and to the senses of touch, taste, smell, sight and hearing. Some students live a normal life until a congenital illness begins to affect them at an advanced age. Muscular dystrophy and spinal muscular atrophy are examples where the student exhibits normal abilities at first but develops increasing weakness in various muscle groups as time goes by.

The psychological effects of the student's disability, particularly those of late onset, may be a major hurdle preventing the student reaching full potential and entering the wider community. The following three points illustrate effects due to the student's primary disability:-

(a) The student may be depressed, and have a decreased self esteem and expectation due to 'being different' and due to the decreased range of opportunity.

(b) The student's personality may fail to develop normally due to a failure to interact socially at various ages (6). Several researchers see the need for young "non-walking" children to be provided with equipment to allow them to stand and/or move about so that they may interact with other children in the playground environment.

(c) The time taken to overcome one aspect of the disability may reduce the time available to develop other skills so that social and psychological development may be stunted (6).

Regency Park generally caters for approximately 200 students. A number of these students attend other schools for gradually increasing periods of time. A recent study by Ahrens (AH1) of the communication needs of the students attending Regency Park revealed that:-

(a) 2% of students had legible handwriting but unintelligible speech,

(b) 7% of students had intelligible speech but illegible handwriting, and

(c) 16% of students had unintelligible speech and illegible handwriting.

2.4.2 Existing Methods

The broad range and degree of disability of students at Regency Park has been discussed above. Some of these students are assisted by "technical aids" provided by the Rehabilitation Engineering Department (66). These technical aids may be divided into three categories (40); 1. orthotic aids, 2. adaptive aids and 3. aids to access general consumer products.(67) An orthotic aid refers to apparatus applied to the body to replace or impose a function or to support, activate or protect parts of the body. The adaptive aid relates to those devices specially developed for the disabled population to allow optimisation in activities of daily living (e.g. mobility aids, communication aids etc.). The third category refers to "interface" devices which allow the disabled individual to access ordinary consumer objects.

In all examples of technical aids categorised as two or three, the technical aid contains or requires a separate interface of some kind that provides physical access to the aid. In this instance, the interface is defined to be a means of access for an individual to an object, mechanism, machine, or system of machines the individual wishes to exploit in the process of carrying out a specific activity.

The interface component is of primary importance because it has considerable influence on user motivation, function, achievement and performance. The interface to a great extent prescribes the effectiveness of the outcome of a user/device relationship. The most serious problem facing both users and prescribers is the lack of knowledge of this intimate linkage between the user and the technical aid (40).

The interface senses some physical action of the user which can be controlled in a conscious and repeatable fashion by the user. For many able bodied people the keyboard is a satisfactory interface for the entry of characters to a computer since they are able to control the motion of their fingers consciously in a repeatable fashion. For many disabled people the keyboard is a very poor interface since their hand movements are too limited, or weak or are affected by spasticity. An alternative interface is required for these people to allow them to take advantage of equipment available to the able bodied.

Interfaces employed in communication devices fall into three categories (55), 1. Direct selection, 2. Scanning and 3. Encoding. Direct selection refers to interfaces where the user specifies the desired character or symbol with a simple direct action which points to or otherwise defines the character. The following are examples of this technique:-

(a) Pointing with a finger or other limb at particular characters arranged in a convenient fashion.
(b) Depressing keys on a keyboard to select characters.
(c) Staring at a character in a "see through" matrix until the therapist guesses the selected character.
(d) Eye gaze devices which monitor fixation of the eyeball (63) or the associated EEG signal (71,72).
(e) Head or limb mounted pointers (66) or optical pointers (56) for selection of characters mounted on physical or photosensitive switches.

Scanning refers to interfaces which step through the set of characters or symbols until a specific signal is received from the user which indicates the currently displayed character is the desired selection. The following are examples of this technique:-

(a) The therapist points to, or says, characters until the user nods or otherwise indicates selection.
(b) A light moves through the range of characters until a switch is depressed. The switch may be activated by any muscle group, e.g. eyebrow switch, foot movement (51), eye movement, pneumatic pressure from the lungs, or by

monitoring the electric signal associated with muscle function (14).

(c) Computer hardware (e.g. the "Adaptive Firmware Card" [1,65]), and software (e.g. Macapple [68]) offer several scanning techniques where characters are displayed as lists or matrices with cursors of varying type or speed which attempt to offer flexibility so that "user satisfaction" may be maximised.

Encoding refers to interfaces which depend on a coded signal from the user. The following are examples of this type of interface:-

(a) The user enters a series of remembered icons on a keyboard which causes the device to utter a predetermined sentence. (e.g. MINSPEAK [68])

(b) The user enters the number of the row and column of the character in a matrix of possible alternatives.

(c) Morse code and other similar bi and tri state codes may be employed by using various muscle groups, including the lungs, to activate switches.

The very wide range of interface strategies is necessary to provide the best interface for the particular user who will have a unique combination of cognitive and physical abilities. It should be noted that all strategies have negative aspects. If the user perceives that the interface is too slow, frustrating, fatiguing, bulky, unattractive or otherwise unacceptable then the aid will not be effective.

Many interfaces employ the skeleton and muscle groups in unusual fashions (e.g. head mounted pointers for typing) which may lead to problems, such as repetitive strain injury, in the long term.

A recent survey at Regency Park (2) found that 10 out of 39 students with unintelligible speech had no communication device. Ahrens (2) noted that there is no suitable communication device for 5 of these 10 students. In many cases where an interface has been provided, the user is faced with a very slow and frustrating task. For example, scanning methods requiring two user inputs per character often yield less than ten characters per minute. At this rate it would take approximately 180 minutes, or 3 hours, of continuous effort to type one page of information.

It may be seen therefore that although there are a large number of interfaces available, there is still a need for further investigation to find faster, more efficient, more portable and cheaper interfaces. There is also a need to quantify what may be expected from the interfaces so that therapists may prescribe the correct device for their clients.

The selection of the patient-device interface is often intuitive. Normally a therapist, or panel of health professionals, select a modality based on their experience of the available interfaces and of the client. This is generally not an easy task because of the diversity in performance of modalities and abilities of clients. The following sets out a method of quantifying the performance of an interface suggested by the author. The literature review, and experimental data associated with this thesis will be used to define the performance of the VIM voice input device as a data input interface for the disabled. It should be noted that the following methodology is proposed by the author. The suitability of the procedure for comparisons of modalities has not been tested.

The suitability of a modality is a function of three characteristics, i.e. Accuracy, Power and Convenience:-

(a) **Accuracy.** Accuracy is a consideration of the error rate, and the presentation of the product. Each application has a required accuracy, and each combination of client and modality will tend to produce a particular level of accuracy. For example, when considering an application an error rate of, say, 5% may be suitable for personal notes or for correspondence to a sympathetic reader. On the other hand a handwritten letter free of errors may be unsuitable as a business letter because of poor presentation due to spastic hand movement. Section 2.4.3. attempts to quantify minimum acceptable accuracies for particular applications at Regency Park.

Similarly when considering the client and proposed modality, thought must be given to the accuracy inherent in the combination. For example a person with spastic hand movement will tend to produce a percentage of errors when using a keyboard which will depend on the severity of the spasticity and the layout and design of the keyboard.

When quantifying the performance of an interface it would be necessary to predict the percentage of attempts resulting in correct entries. It may be necessary to define a range as a function of the user's disability.

(b) **Power**. This characteristic is a function of the speed of data entry and the vocabulary size. Consideration of Chinese and European, key operated typesetting machines illustrates the problem. In the European machine several single characters are rapidly selected from a relatively small set of characters to produce a word. In the Chinese machine the operator slowly selects from a very large set of characters, but the characters represent a whole word so that the overall "user satisfaction" may be similar for both processes.

The problem of comparing and selecting interfaces for the disabled is more difficult because of the variations due to the techniques of direct selection, scanning and encoding and due to the various parts of the body which may generate the signal. Norman (49) introduces the concept of a psychological measure of "user satisfaction". He suggests that we may determine the impact of changing physical parameters upon the psychological variable of user satisfaction. We can then objectively assess the desirability of tradeoffs among the parameters.

Norman defines a function U(x), the user satisfaction for attribute x, and defines its form to be $U(x) = kx^P$. Norman points out that a typical result of psychological scaling is that a substantial increase in the current value is required to make the increase worthwhile. He illustrates this by referring to menu size. He suggest that if the user satisfaction of a menu of 300 is given a value of 50, i.e. U(300) = 50, then the menu would have to increase to 1000 to double the user satisfaction, i.e. U(1000) = 100. Solving the equations implies, in this example, that $U(S) = 1.9 \times S^{0.6}$.

In a similar way equations are defined below for user satisfaction for vocabulary size (V) and speed of data entry (R). It should be stressed that values in the following equations have been arbitrarily selected to demonstrate the methodology.

(i) If the user satisfaction associated with a vocabulary of 100 items is 1 and the vocabulary must be increased to 400 to double the satisfaction then:

$$U(V) = .1\ V^{0.5} \qquad \text{Equation 1.}$$

(ii) If the user satisfaction associated with a speed of data entry of 100 entries per minute is 1 and the speed must be increased to 400 to double the satisfaction then:

$$U(R) = .1\ R^{0.5} \qquad \text{Equation 2.}$$

The performance characteristic "Power" is a measure of the expected user satisfaction due to the speed and vocabulary size of the interface. Here we define "Power" to be the product of equation 1 and 2.

$$P = .01\ V^{0.5}\ R^{0.5}$$

To illustrate this concept further consider the following examples:

(i) Figure 3 uses the data from the indicated references to plot the power of voice input devices.

(ii) A typist entering 20 words per minute (assume 5 characters per word) using a keyboard with approximately 100 keys (considering upper and lower case) has the following power:-

$$P = .01 \times (100)^{0.5} \times (20x5)^{0.5}$$

$$P = 1.$$

If the typing rate increases to 100 w.p.m. then P = 2.24.

(iii) A person using a scanning technique to pick words from a 15 x 15 array takes two inputs to select each word. For the purpose of this example assume that each input takes, on average 12 seconds.

$$P = .01 \times (15x15)^{0.5} \times (60/(12x2))^{0.5}$$

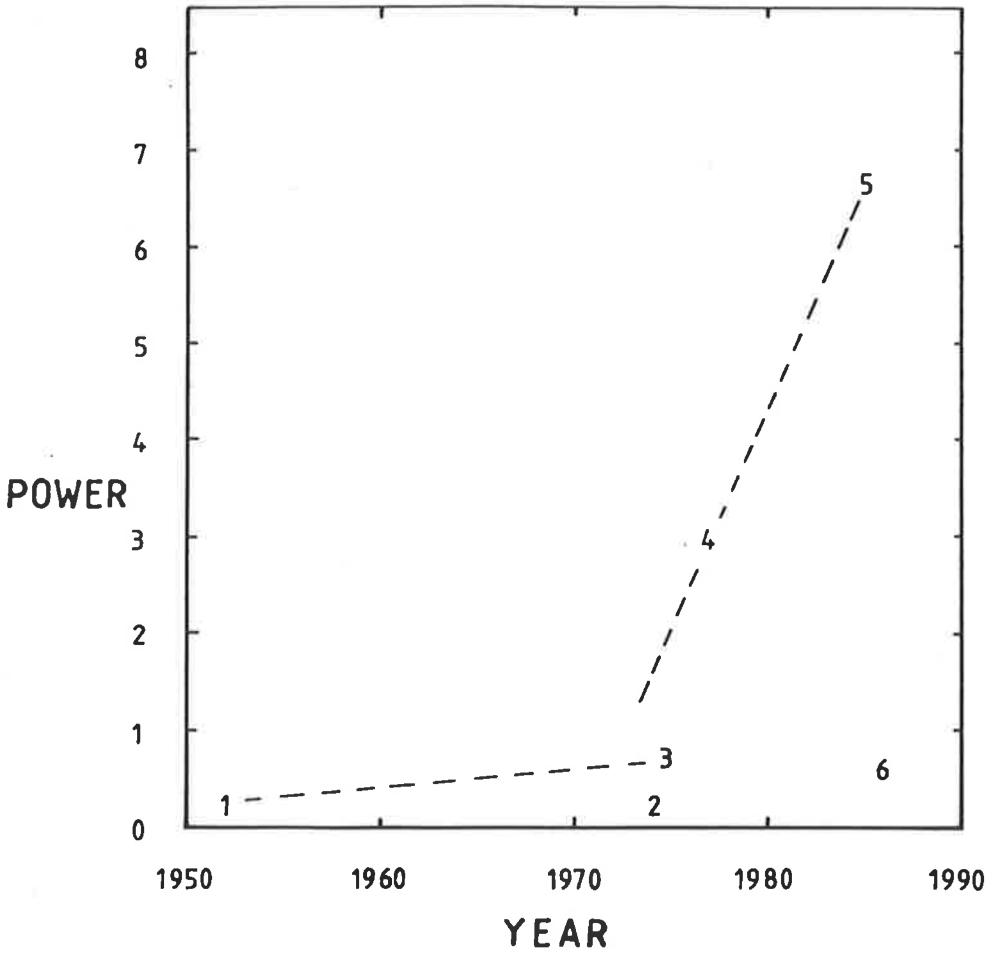
$$P = 0.24.$$

(iv) A person uses a head mounted torch to select items off a 10 x 10 matrix. The person takes, on average, two seconds to select the item and must fixate for three seconds to activate the item.

$$P = .01 \times (10 \times 10)^{0.5} \times (60/(2+3))^{0.5}$$

$$P = 0.35.$$

(c) **Convenience.** "Convenience" is a summary of factors peculiar to the device which should be taken into account. Items which might be noted include:- cost of the device, portability, tendancy to produce muscular pain and difficulty in setting up the device.



NO.	REFERENCE	YEAR	RATE C.P.M.	VOCAB. SIZE	POWER	ACCURACY + COMMENTS
1	(20)	1952	15*	10	·12	98 %
2	(42)Martin	1974	30	34	·32	98·5%
3	(42)Itakura	1975	25*	200	·71	97·3%. 22 TIMES REAL TIME.
4	(37)	1977	90*	1011	3·0	95%. 80 TIMES REAL TIME.
5	(33)	1985	90	5000	6·7	95%
6	This Thesis	1986	34	80	·52	95%

*RATE NOT GIVEN IN REFERENCE, ASSUMED FROM DESCRIPTION.

Figure 3. The Development of the "Power" of voice input equipment.

Power = $.01\ V^{0.5}\ R^{0.5}$

where:- R = Characters per minute

V = Vocabulary size.

2.4.3 Potential of Voice Input at Regency Park

This subsection considers the potential offered by voice input technology for the disabled community served by Regency Park and identifies areas for development. "Potential" here means a consideration of the advantages which might be gained, and the number of people who might gain them, compared with an assessment of the expected performance and whether development at Regency Park is appropriate.

Data Entry

Data Entry is very important for the students at Regency Park both in the short and long term. Normal school curricula require students to write when recording new material, learning the material and when preparing material to show that the subject has been mastered. In the longer term disabled people with mobility problems naturally tend towards sedentary occupations. These occupations often require the preparation of large amounts of written or typed material. Clearly, therefore, efficient methods of data entry are required. In the future the technology currently being developed (29,33) will allow people without hand function but with good voice control to type as quickly as able bodied typists. In the short term, however, voice input is a relatively slow and laborious means of data entry. Even so, for some it is the best method available (ref. 2.3.2).

When considering voice as a modality for data entry, it is important to consider the required accuracy of the particular application. Some applications, for example the entry of computer code, require 100% accuracy. In other applications an accuracy less than 100% may be acceptable. Students preparing material for their own future reference, or when preparing material for a sympathetic reader may find a lower accuracy acceptable. Appendix A6-C shows a paragraph typed with 100%, 95%, 90%, 85% and 80% accuracy. It is suggested that the paragraph typed with 95% accuracy may be easily understood. The paragraph with 90% accuracy is more difficult to read but may still be understood. The paragraphs with 85% and 80% accuracy are difficult to understand. Some games and arithmetic practice programs, for example Wuzzle and Spaceship, handle incorrect entries by showing the correct answer. In these circumstances an accuracy of say, 80% may still be acceptable. If it is reasonable to expect the user to improve the accuracy of the final material by 5% by correcting the data then the following accuracys may be taken as minimum required values:-

Description	Minimum Accuracy.
(i) Demanding applications (eg computer code)	95%
(ii) General applications (eg letters, essays)	85%
(iii) Error tolerant applications (eg games)	80%

In summary, there is a significant proportion of the population served by Regency Park with an unsatisfied need for data entry devices. Inexpensive voice recognisers have been shown to meet this need elsewhere, therefore there would appear to be strong potential for voice input for data entry at Regency Park.

Environmental Control

Caudrey's (13) analysis of the technological priorities for enhancing integration of Regency Park students found that the areas of greatest perceived need (listed in order of priority) are:-

(i) mobility

(ii) toileting

(iii) activities of daily living

(iv) speech

(v) injury risk.

The first three of these, in particular, would seem to be amenable to solutions involving technical aids. These aids would then become a part of the student's environment and would be controlled by the student's most appropriate ability.

Wheelchairs are a very significant mobility aid for many students at Regency Park however voice control has, to date, not been shown to be a practical control method. Clearly progress in this field should be monitored to take advantage of advances in technology.

The suitability of voice control of other technical aids would have to be taken on a case by case basis. In general the technical aids to meet the above needs are not yet developed. It is therefore premature to consider voice input at this stage, although the developers should be aware that voice input might be used. Most documented voice controlled environmental control units are located in the user's bedroom or home due to the nature of the controlled devices. Regency Park's main thrust is to integrate its students into society so only limited work has been done in this area. Environmental control by voice input, therefore has little potential at Regency Park in the short term. However it is desirable and appropriate to quantify each student's abilities with voice control and compare them with that student's other control modalities. This will assist therapists to provide suitable technical aids for the student in the future. It may be seen that this is really an assessment of the student's ability to use voice for data input and therefore may be considered within that topic.

Speech Therapy

Ahrens (2) reported that 18% of students served by Regency Park have unintelligible speech. Clearly there is a very strong need for techniques which will overcome this situation. A Speech Therapy department is located within Regency Park and works with the students to overcome communication problems. Research into the use of voice recognition equipment in this area is still at an early stage, however available reports indicate positive results. If equipment does prove to be of benefit then it will have strong potential at Regency Park.

Aids For The Deaf

Regency Park does not specialise in providing strategies and technology for clients to overcome deafness. This work is performed by other institutions so there is little potential for the development of voice recognition equipment to aid the deaf at Regency Park.

Translation

Devices similar to Fright's (25) speech translator could be used by a number of students at Regency Park. However development of such equipment is at such an early stage that it is difficult to assess the potential.

The technique reported by Schmitt (64) is a combination of voice input to enter a string of data, and then a synthesizer to enunciate the string. This concept is quite similar to several highly developed devices available in the market place (80) which do not use voice for data entry. Voice input should be considered for students who will use a speech synthesizer, however the question would be one of deciding the most appropriate modality for data input.

Considering the current performance of this type of technical aid there is little potential for use at Regency Park. With the current state of equipment the primary question is whether voice input is the most appropriate data entry technique for the given person.

Summary - Development Directions at Regency Park

In summary, the two areas requiring development work in voice input at Regency Park are:-

(i) To assess and develop voice input as a data entry technique.

(ii) To assess and develop voice input as a tool for Speech therapists.

The project methodology developed in chapter three centres on these two issues.

3. PROJECT METHODOLOGY

3.1 Introduction

The preceding sections have illustrated the widespread belief that voice input will be of benefit to the disabled. Advances in speech input technology have generally been gradual as the various manufacturers have applied innovations. It is unlikely that there will be a sudden change which will make speech technology suitable for, and available to the overall disabled population. It is more likely that the equipment will gradually become more effective, portable and inexpensive so that it will gradually become suitable for an increasing section of the disabled population.

Several cases (section 2.3) have shown that speech input is suitable for a small group of disabled people. Generally the studies report success in two situations:-

(a) Situation one. Subjects characterised by very strong motivation, reasonable voice control, good cognitive skills and gross disability disallowing most other interfaces except bi and tri state switches, or more exotic modalities such as eye gaze.

(b) Situation two. Therapists reporting improved speech, or attitude when using speech input.

In the light of the above reports the Rehabilitation Engineering Department at Regency Park purchased a Voice Input Module (VIM) in 1984 to investigate the suitability of the modality for students at Regency Park. Selection of the device was based on:-

(a) A recommendation from staff at the Children's Hospital at Stanford.

(b) An appraisal of the specification which suggested that it was equal to or better than similarly priced recognisers available at that time.

(c) Positive results obtained with the device by Fried Oken (23).

(d) The cost of the device; approx A$1700, which was considered to be an amount which could be afforded if the device offered significant potential.

A study of the population served by Regency Park was undertaken by Ahrens (2) which assessed the assisted communication needs in general terms, and identified a small group of students who would gain from an effective speech input device.

This thesis reports on the experimental program set up to evaluate the selected voice input device. The work has been carried out in three distinct phases:-

(a) Phase one was an initial familiarisation with the device to confirm manufacturers data, and to develop hypotheses to be tested and suitable experimental strategies. This initial familiarisation combined with an appreciation of the needs of the students at Regency Park, and reference to published work on the subject led to the identification of two areas for further work. These areas became the subject of phases two and three.

(b) Phase two tested hypotheses designed to evaluate the potential of the device as a data input modality for disabled students.

(c) Phase three tested hypotheses designed to evaluate the potential of the device as a tool for speech therapists employing repetitive articulation drills to improve their clients' articulation.

3.2 Hypotheses to be Tested

The primary hypothesis and six secondary hypotheses tested during this thesis are listed below:-

(a) Primary hypothesis:

That the VIM sound recogniser is of practical use to members of the community served by Regency Park.

(b) Secondary Hypotheses:

(i) Hypothesis tested in phase one:-

H1 That the VIM may be used by an able bodied speaker to enter data to a computer with an accuracy which would be suitable for use in applications at Regency Park.

H2 That the VIM may be used by an able bodied speaker simulating dysarthric, but repeatable, sounds to enter data to a computer with an accuracy which would be suitable for use in applications at Regency Park.

(ii) Hypotheses tested in Phase two:-

H3 That the VIM may be used by a trial group of disabled children for data entry into a computer.

H4 That the VIM is the preferred modality for a trial group of disabled children for data entry in a classroom environment.

(iii) Hypotheses tested in Phase three:-

H5 That the VIM may reliably discriminate between sounds representing common articulation errors when the VIM compares the speakers sounds with reference patterns produced by someone other than the speaker.

H6. That the VIM may reliably discriminate between sounds representing common articulation errors when the VIM compares the speakers sounds with reference patterns produced by the speaker.

A relatively small group of disabled people are benefiting from speech input devices. This thesis primarily tests the hypothesis that recent advances in technology have made voice input suitable for a broader section of the disabled community, in particular, the people served by the Regency Park Centre for Young Disabled.

Poulton (53) stated that "An aid will be used if the disadvantages inherent in its employment are outweighed by the benefits derived by the user." This philosophy has been the underlying theme for the test of the primary hypothesis. If after appropriate steps have been taken to implement the equipment we find that it is being used voluntarily because its benefits outweigh its inherent disadvantages then the hypothesis will be shown to be correct. A methodical strategy has been used to take the device out of the laboratory and into the normal Regency Park environment. The work has been performed within the Rehabilitation Engineering Department at Regency Park with the author being responsible for the assessment and enhancement of the equipment and the support of clinical staff during initial use and testing. The trials of the device have been conducted by, or have involved teachers and therapists in typical clinical settings. The trials have also been run over a significant period of time with a large number of participants, both therapists and clients. This has tended to reduce short term effects such as short term obstacles, and enthusiasm of individuals. For these reasons, the degree of acceptance of the device in the long term is a reliable test of the hypothesis.

In testing the primary hypothesis, six secondary hypotheses have been identified and tested. These secondary tests have been selected to be milestones in the project. The results of the tests have been used to modify the methods used in succeeding work. The six secondary hypotheses (H1 to H6) are associated with the three phases of the project described in the introduction to this chapter.

3.3 Method

This section discusses in general terms the methods, and clinical relationships during the three phases of the project. A detailed description of the test procedures is in chapter five.

(a) Effort was initially directed at developing an understanding of the VIM sound recogniser.

Methods used included:-

(i) Familiarisation with the supplied manual and software.

(ii) Interfacing routines between the Apple computer and the VIM device were established by analysing and flowcharting the supplied software.

(iii) Reference to published work provided an insight into the principles of operation of the VIM and the possible potential and limitations.

(iv) Estimates of the performance of the equipment, through informal experiments.

The knowledge of the equipment gained during the familiarisation period was used to develop the hypotheses listed above. Hypotheses H1 and H2 were then tested. The experiments testing these hypothesis were conducted in a laboratory environment with the author's voice. It was noted during this stage that the programs supporting the VIM might be modified to be more useful for teachers and therapists. Development of a computer program to be used in a classroom environment was therefore started. The development was primarily aimed at minimising unnecessary options and file manipulation to minimise workload and distraction in the classroom. This is discussed in more detail in Chapter four.

(b) The positive results obtained during phase one led onto testing of the VIM and the new software. A small group of children who appeared to be potential users of voice input were selected and introduced to the equipment, initially in the electronics laboratory. Initially training and testing was carried out by the author. As the work progressed a special education teacher became involved. Following successful testing of the new software, and after initial testing of the children's performance, the VIM was moved into the computer resources classroom. Further training and testing of the children's performance continued in the classroom supervised by the special education teacher. Data collected throughout this period was used to test hypotheses H4 and H5.

(c) The possibility of using the VIM as an articulation training aid was identified during phase one. A simple program was developed which was used to illustrate the application to the speech therapy department. A specification for a program which might be useful in the clinical situation was developed. Ease of use, appropriate feedback, and good discrimination between good and bad articulation were identified as primary goals. The VIM requires that examples of good and bad articulation be pre-stored in the computer's memory for reference. The therapists indicated that it would be desirable to use a therapist's speech for the reference sounds because of the difficulty experienced by clients in training examples of good and bad articulation into the VIM. The importance of the VIM's discrimination performance led to the selection of hypotheses H5 and H6.

Experiments designed to test hypotheses H5 and H6 were carried out with able bodied, cooperative adults saying words selected to represent common articulation errors. Following testing it was felt that the equipment did not show sufficient promise to justify trials with dysarthric children. It was felt that the anticipated poor performance of the equipment could be misconstrued by the children as a fault on their part leading to decreased self esteem.

4. SOFTWARE DEVELOPMENT

4.1 Introduction

Software has been modified and created to make the voice input equipment more suitable for the applications considered. This chapter explains why it was necessary to develop the supplied software and explains the goals and rationale behind the work. The two major programs developed during the project are discussed in general terms in this chapter. The detailed operation of the programs is illustrated in the User's Manual in Appendices A2 and A4. The programs are also provided for inspection in Appendices A3 and A5 and on the floppy disc in the rear pocket of the thesis.

Several smaller programs were also developed during the project. They are mentioned briefly here to illustrate the extent of work undertaken.

(a) During initial testing (phase one) a program was developed which prompted sounds in random order. The program did not display the results of the VIM's analysis, but saved the results for printing at the end of the session. The program employed these features to avoid influencing the subject's behaviour during the experiment. Appendix A6-a is a copy of the output of this program.

(b) During initial testing short programs were created to informally test the performance of the VIM as an environmental control unit.

(c) A rudimentary articulation training program was developed during initial testing. This program was used to demonstrate the basic concepts to therapists who were then able to help develop a specification for the final articulation training program.

(d) Several short programs were created to assist in the generation of graphics files for the Voice Tutor Program.

4.2 VIM Background

The VIM voice recogniser used during this project is a speaker dependent isolated word recogniser. Physically, the VIM is a circuit board which is mounted into one of the expansion slots in the Apple II computer. The VIM circuit board contains input signal conditioning circuitry, a speech preprocessor, an analog-to-digital converter, a recogniser/controller which performs the normalisation and classification functions and memory for storage of programs and reference voice patterns. The circuit board handles all tasks associated with analysis of the speech signal leaving the host computer free to run standard software. However, at certain times, the host computer must run a control program to set parameters on the VIM circuit board, set the reject threshold, and to select recognition or training modes. The control program interactively coordinates the functions performed by the VIM circuit board with the requirements of the user. For example, during training the control program prompts the user through the required steps, automatically setting the circuit board as required.

The control program supplied with the VIM is called AVIM. AVIM is a menu based program written in Applesoft Basic which allows the user to manipulate all the functions available on the VIM. AVIM may be used to:-

(a) Build and edit vocabularies.

(b) Train sounds into the VIM (i.e. set up reference patterns).

(c) Set the VIM into recognition mode.

(d) Test recognition performance.

(e) Set and read parameters.

While AVIM is not difficult to use it does assume a knowledge of the use of computers, and requires some time to become familiar with the various options and techniques. File names must also be used when saving or retrieving voice patterns.

In summary, AVIM is designed to be a powerful utility program able to fully utilise the capabilities of the VIM circuit board. However the cost of this versatility is that the operator must be familiar with the AVIM program and its functions. To run an application the user may have to use several menus and options entering and manipulating files at various times. While this is not difficult, it does take some time.

4.3 Teacher's VIM

During the initial testing it became clear that a program like AVIM would be required in the classroom to train and test new voice patterns and to set and activate the VIM circuit board. However it was also clear that AVIM was less than ideal because:-

(a) of the time required to learn to use AVIM.

(b) of the additional workload and distraction in the classroom due to working through various menus and manipulating files. Several early sessions were disrupted when the operation of the AVIM program distracted the student and operator away from the main purpose of the session.

The following specification was developed for a suitable program:-

(a) The program should start automatically when the power to the computer is turned on.

(b) The program should default to load voice patterns, activate the VIM and run the application program unless the user intervenes. This is in recognition of the fact that training should only be required when first using the VIM. This feature would also be desirable in the long term where the computer would be turned on by, for example, the weight of the user's wheelchair. This would allow operation of the computer with voice alone.

(c) The program should be specific to the classroom application requiring the minimum of effort and distraction.

(d) The program should be consistent with AVIM and should allow the user to transfer to AVIM if required.

These goals were achieved by modifying the AVIM program to produce the program Teacher's VIM which is shown in the flowsheet in Figure 4. The operation of the program is explained in detail in the User's Manual in Appendix A2. AVIM is modified in two areas:-

(a) The beginning is changed to allow a default which loads and activates the VIM and runs the application program.

(b) An extra menu is added. Care has been taken to select only the options required by a teacher, to put them in the order of use and to supply sufficient prompts.

4.4 Voice Tutor

Initially an application program was developed which allowed a demonstration of the VIM in an articulation training role. The subsequent comments from the speech therapist led to the following specification:-

(a) The equipment should discriminate between two to five sounds. The number of sounds and the voice patterns in memory should be readily changed during the session.

(b) The reference sound patterns should be readily saved to and loaded from disc.

(c) That a changing image on the screen of the monitor would be an

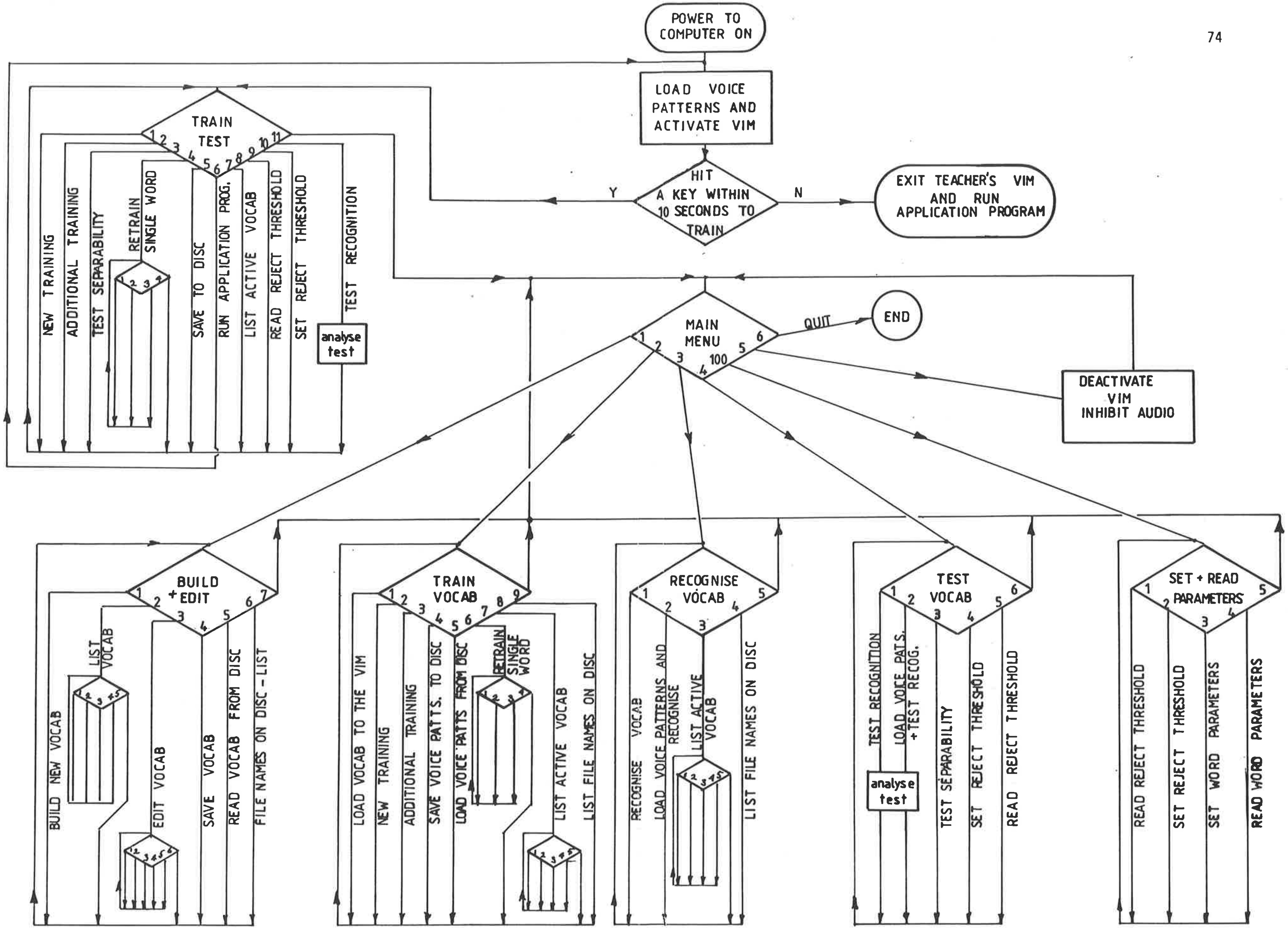


FIGURE 4. Teacher's Vim Flowsheet.

appropriate feedback mechanism. Both positive and negative feedback in visual and aural form would be desirable, however it should be possible to easily suppress these options to suite the particular client.

(d) That the reject threshold should be easily changed at any time during the session.

(e) That the display should readily show the therapist the status of the various options, and the results of the VIM's analysis. For example, if the VIM found that the uttered sound was not correct then the therapist would wish to know which sound the VIM found was closest.

(f) That control and monitoring activities by the therapist should cause the minimum interruption to the session.

It may be seen that the requirements set out above are fundamentally different to the characteristics of the AVIM and Teacher's VIM programs already discussed. The programs already discussed set up the VIM and then transferred to the application program. The user of the Voice Tutor program however would wish to be able to control and reset the VIM at any time while running the application program.

These goals were achieved by:-

(a) Marrying the VIM control functions and the application program into the one program called Voice Tutor. Control commands and application program commands can therefore be made at any time.

(b) Providing all information for both the therapist and the client on the one screen. The bottom four lines of the screen detail the status of all options and the results of the VIM's analysis for inspection by the therapist. The use of any option by the therapist, except the help screen, does not influence the client's display which minimises interruption to the session.

Appendix A4 is the Voice Tutor Users Manual which details the operation of the program. Figure 5 shows the flowsheet for Voice Tutor.

Voice Tutor has not been used in a clinical setting for the reasons discussed later in this thesis. However it proved invaluable in the testing of the hypotheses related to articulation training because the VIM's settings could be easily altered, and the analysis made by the VIM could be easily seen and recorded.

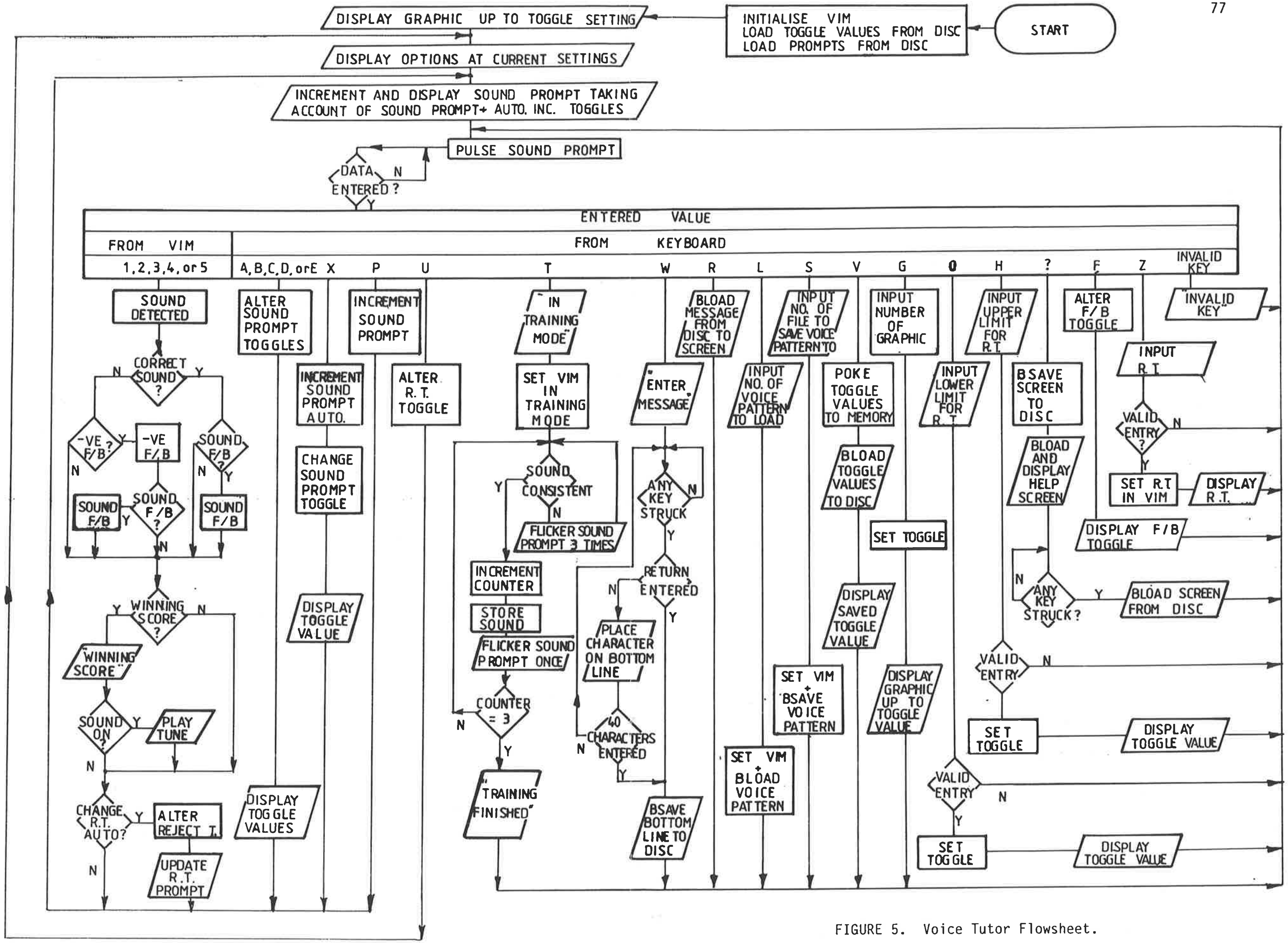


FIGURE 5. Voice Tutor Flowsheet.

5. TESTING

5.1 Introduction

This chapter discusses in detail the experimental work which has been done to test the secondary hypotheses described in chapter 3. The data is interpreted and conclusions are drawn on the hypotheses taken in isolation. Chapter 6 discusses the primary hypothesis and takes an overview of the interpretations presented in this chapter.

The chapter is divided into the three phases described earlier and looks at the experimental work under the following headings:-

(a) Hypotheses tested

(b) Subjects

(c) Design

(d) Apparatus

(e) Procedure

(f) Results

(g) Interpretation.

Phases 1 and 3 employed able bodied volunteer speakers. The testing programs tended to be rigorous involving repetition and discipline to provide a statistically relevant database. An estimate of the confidence level of derived statistics is shown where possible and appropriate.

Phase 2 involved subjects drawn from the student body at

Regency Park. In this case the prime concern was the well being of the students so that a rigorous test procedure was not adopted. The students were trained on the speech input device until the teacher felt that their performance had stabilised. Their performance was then assessed by monitoring normal classroom activities.

5.2 Phase One - Initial Testing

(a) Hypotheses tested:-

H1. That the VIM may be used by an able bodied speaker to enter data into a computer with an accuracy which would be suitable for use in applications at Regency Park.

H2. That the VIM may be used by an able bodied speaker simulating dysarthric, but repeatable, sounds to enter data to a computer with an accuracy which would be suitable for use in applications at Regency Park.

(b) Subject

The author's voice was used throughout the testing of hypotheses H1 and H2. The author was an able bodied 29 year old male at the time of testing. His native language is English with an Australian accent.

(c) Design

The experiment quantifies the accuracy of the VIM when used by the subject to enter data. The subject was experienced in the use of the VIM and adopted the techniques recommended in the user's manual (76). The following variables were altered during the experiment:-

(i) The reject threshold was varied throughout its range to find whether accuracy varies significantly with reject threshold, and to determine the accuracy of the VIM at the optimum reject threshold.

(ii) The subject used three series of 26 sounds to enter data into the computer. The series were:-

1. The alphabet. The subject uttered the sounds normally associated with the letters of the alphabet.
2. The International Civil Aviation Organisation (ICAO) alphabet. The subject spoke the words "alpha", "bravo", "charlie", etc. to enter the letters A, B, and C etc. into the computer.
3. Morse Code. To simulate dysarthric but repeatable speech the subject used combinations of two phonemes "di" and "dah" to enter the letters of the alphabet into the computer. The combinations correspond to the Morse code. Dysarthric speech may take many different forms. This experiment attempts to model a form of dysarthria where the person is able to pronounce two syllables in repeatable patterns.

(d) Apparatus

The VIM was set up in an Apple IIe computer in accordance with the manufacturer's instructions. The subject's voice patterns were entered into the VIM's vocabulary. The three series of 26 sounds gave a total of 78 sounds. The VIM has a capacity for 80 sounds, the other two positions were filled by the words "correction" and "return".

A computer program was developed and entered into the computer. The program performed the following functions:-

(i) Automatically prompted sounds from the three vocabulary series. The program automatically displayed prompts associated with the sounds, for example, "Alpha", "A" and "di dah", on the computer monitor. The prompted sounds were displayed in random order from all three series. This was done to average the effects of fatigue and to prevent the subject anticipating sounds and changing pronunciation accordingly. The program showed all 78 vocabulary items once.

(ii) Following display of the prompt the VIM was activated. When a sound was detected, the VIM issued a unique set of characters to the computer which corresponded to the sound recognised. If the characters received by the computer from the VIM matched those prompted then a correct data entry was achieved.

(iii) After the VIM received the sound and the comparison was made the prompt was removed from the screen without showing the subject whether the data was entered correctly. This was done to prevent the subject reacting to results of comparisons. The results of all 78 comparisons were stored and typed out after the 78th comparison was made. Appendix A6-a shows a typical results sheet.

(e) Procedure

The following procedure was repeated sixty four times during a three month period.

(i) A reject threshold was selected and entered into the VIM. The selected reject threshold was one of fourteen values covering the range of acceptable values. Each reject threshold was used five times during the experimental period. The selection of reject threshold for succeeding test runs was random to minimise effects due to anticipation and fatigue.

(ii) The computer program was activated and prompted sounds from the 78 sound vocabulary in random order.

(iii) The computer automatically recorded the prompted and recognised sounds without feeding back information on accuracy to the subject. The subject recorded, by hand, the number of repeats required due to rejection errors, and spurious recognitions received due to background sounds. If any prompted sound required more than 20 repeats then the entry was taken to be an error and the repeats required was taken to be 20.

(iv) At the end of each 78 sound test session the results were printed out and a rest period taken to avoid effects due to fatigue.

(v) The tests were conducted in a laboratory environment with occasional background noise due to local conversations and noises from adjoining rooms.

(f) Results

The computer printout shown in appendix A6-a is the result of a typical test run. The results of the 64 test runs are shown in the table in Appendix A6-b. The table also shows the mean, standard deviation and 95% confidence interval of the mean of the results. The analysis assumes that the results approximate to a normal distribution, and uses the method for small sample sizes described by Walpole (79).

The data from Appendix A6-b has been used to produce the graphs shown in Figures 6, 7 and 8 on the following pages.

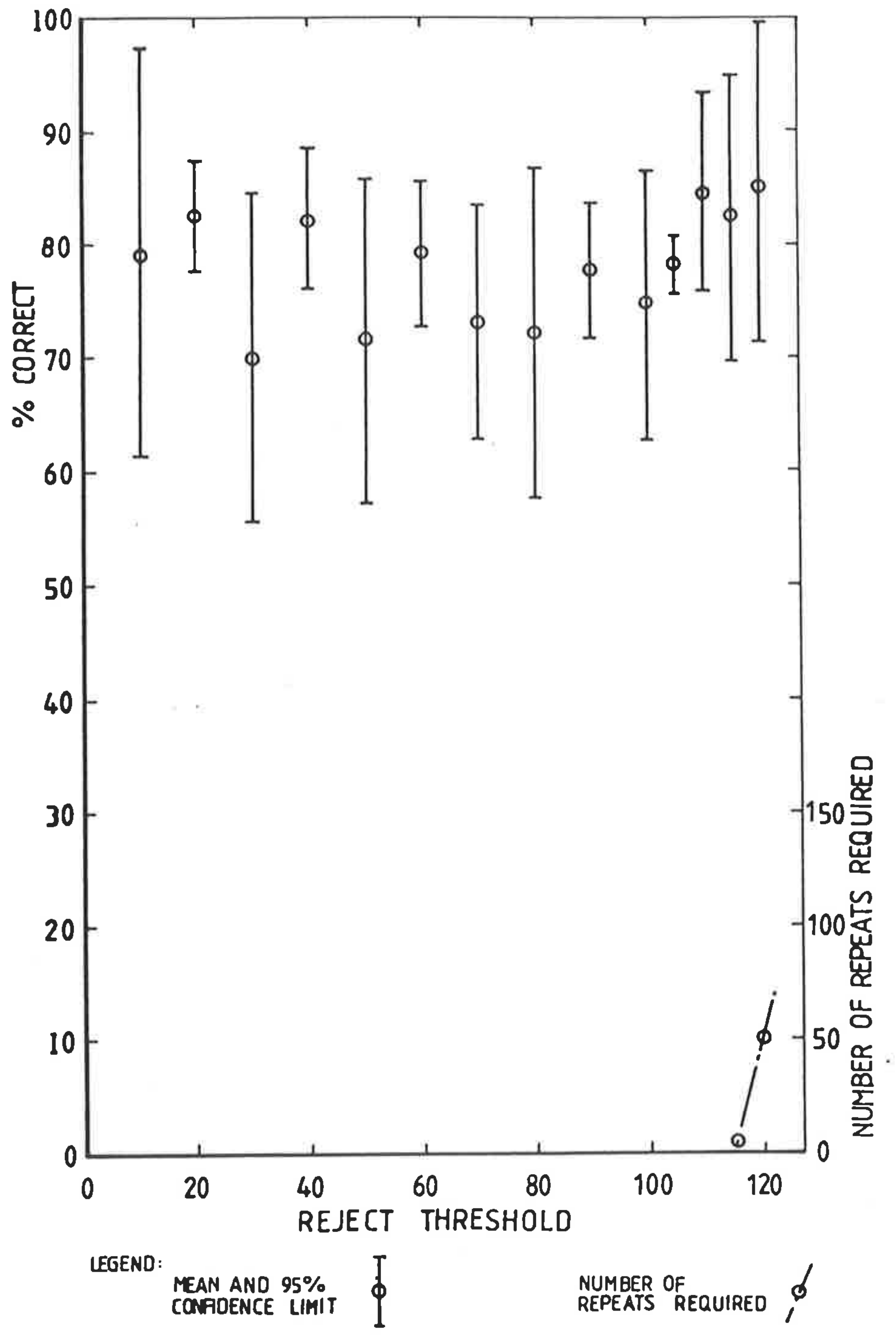


FIGURE 6. The accuracy and number of repeats required when using the sounds normally associated with the alphabet to enter data into a computer, with various reject threshold settings.

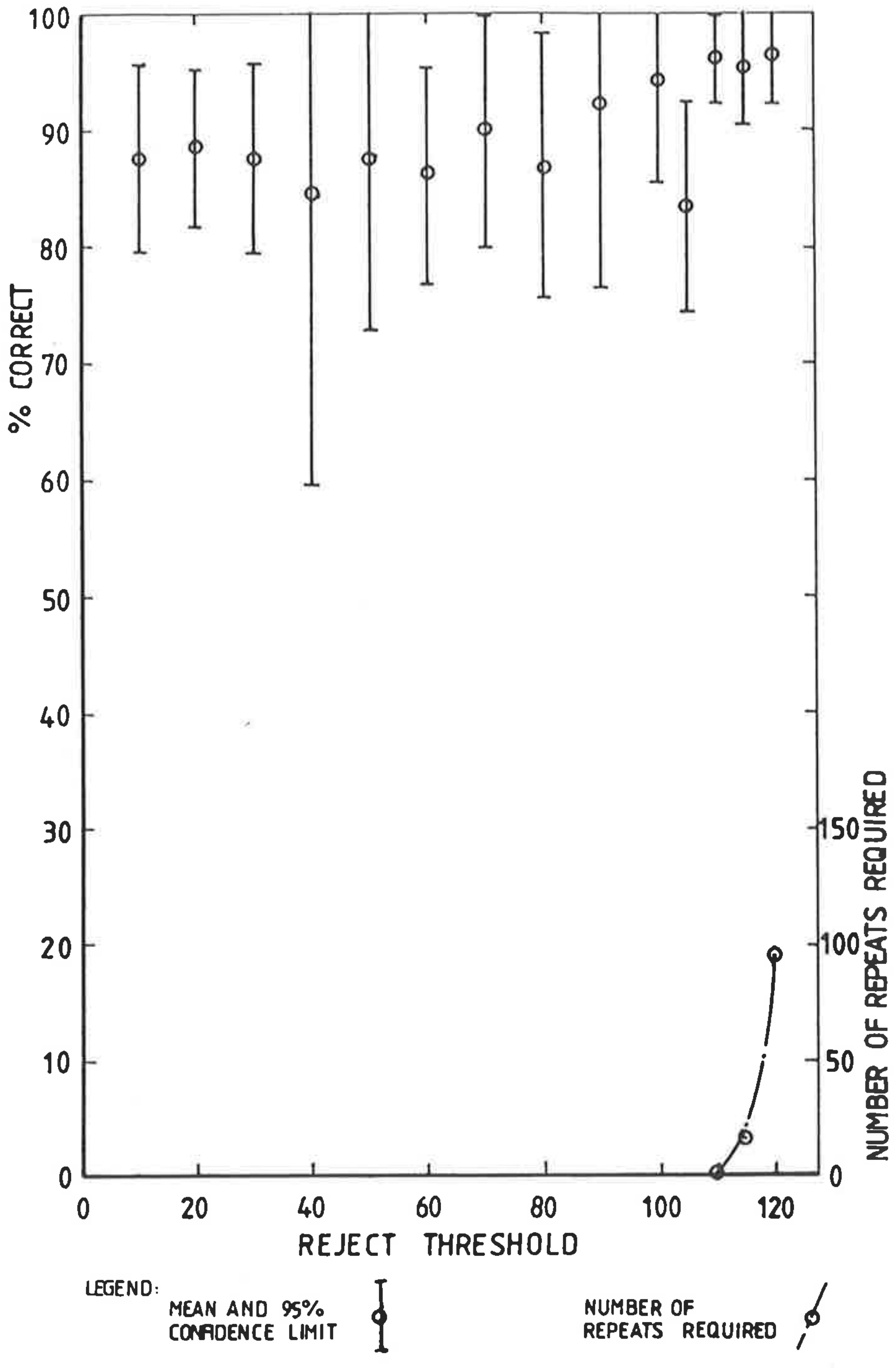


FIGURE 7. The accuracy and number of repeats required when using the I.C.A.O. alphabet to enter data into a computer, with various reject threshold settings.

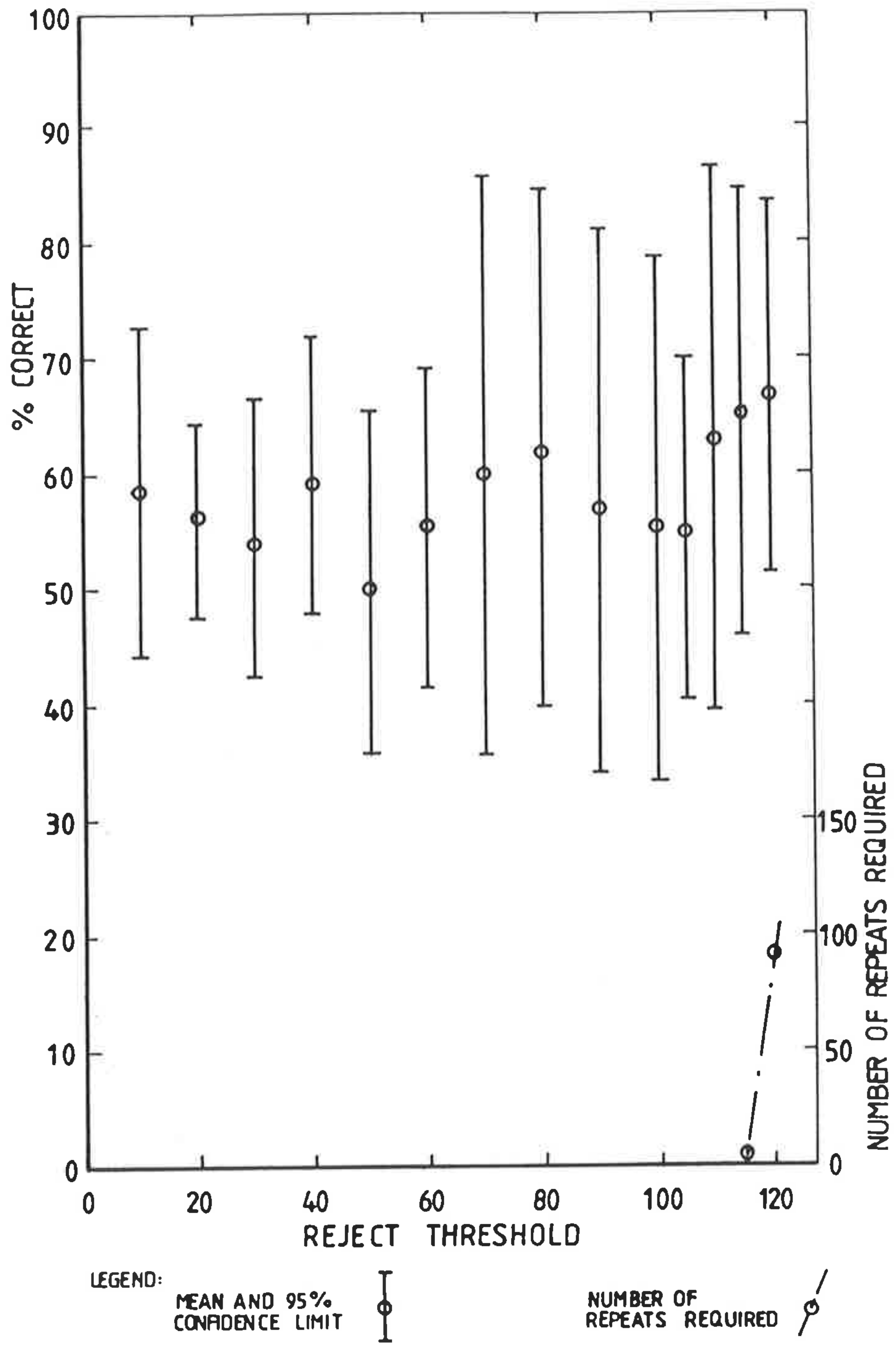


FIGURE 8. The accuracy and number of repeats required when using sounds based on the Morse code to enter data into a computer, with various reject threshold settings.

(g) Interpretation

It is clear from figures 6, 7 and 8 that all three methods may be used to enter data but with significantly different accuracy. Appropriate selection of reject threshold to maximise accuracy while keeping rejection errors to an acceptable level produces the following accuracy:-

	Description of Pronunciation	Accuracy
(i)	Alphabet	82%
(ii)	I.C.A.O.	95%
(iii)	Morse Code	63%

Section 2.4.3 stipulates the minimum required values for accuracy in applications at Regency Park. Comparison shows that:-

(i) The I.C.A.O. style of pronunciation produces data entry accuracy suitable for "general" and "error tolerant applications" but would be borderline for "demanding applications".

(ii) The "alphabet" style of pronunciation produces data entry accuracy with only borderline suitability for error tolerant applications.

(iii) The "Morse code" style of pronunciation produces data entry accuracy which is unsuitable for use.

Therefore hypothesis H1, of page 79, is supported and hypothesis H2 is not supported.

5.3 Data Input

(a) Hypotheses tested:-

H3. That the VIM may be used by a trial group of disabled children for data entry into a computer.

H4. That the VIM is the preferred modality for a trial group of disabled children for data entry in a classroom environment.

(b) Subjects

The voices of nine students were used during the testing of hypotheses H3 and H4. The students varied in age, sex and disability as shown in the table in Figure 9. They were chosen because their current data entry modality is slow or difficult and, it was felt, their speech would be able to activate the VIM. Five of the subjects have cerebral palsy which affects their speech to varying degrees. Subject 3 has virtually normal speech. Subjects 1,2 and 4 are easily understood, however their speech is occasionally interrupted. Subject 7 is very difficult to understand although sounds appear to be fairly repeatable. Subject 6 has easily understood speech however she relies on an artificial respirator. She therefore has reduced control over the timing of her speech, since she must wait for an exhalation cycle, which tends to decrease the rate of her speech. Subjects 5,8 and 9 have clear, easily understood speech.

(c) Design

The experiment involved the subjects in the training and the use of voice input in conditions which might normally be found in the classroom. The performance of the subjects in normal classroom data entry activities was then monitored and quantified. The gathered data was then used to test the hypotheses.

(d) Apparatus

The VIM was set up in an Apple IIe computer in accordance with the manufacturer's instructions. Software, called "Teacher's VIM", was developed from the manufacturer's supplied software and was used to train the subjects in the use of the VIM. A variety of microphones were trialled during the study. Lapel and head mounted microphones were used as well as the table mounted microphone supplied with the equipment. The body mounted microphones were used to overcome problems associated with involuntary body movement.

(e) Procedure

Subjects 1 to 5 were initially trained in a laboratory environment in 30 minute sessions by the author. After approximately five hours of training the special education teacher in charge of computer education took over the role of training the children. The training, which was performed in accordance with Appendix A2, continued in the computer classroom for a further five hours for each student. After this training the subjects' performance appeared to stabilise. At this stage the subjects were encouraged to use the voice input equipment to perform normal classroom tasks such as arithmetic exercises, the writing of short stories and transcription exercises. The speed and accuracy achieved was recorded for later analysis. The training and testing of subject 6 was carried out by the special education teacher. The training in accordance with Appendix A2 was continued for nine hours. Performance tests were then carried out by requesting the subject to transcribe passages from documents using voice input.

Subjects 7 to 9 initially trained for a limited time with the special education teacher however this was discontinued when the teacher received a promotion away from Regency Park. After several months, members of the psychology department took over the training of these subjects using the techniques described in Appendix A2. The sessions were generally carried out in the electronics laboratory. Subject number 8 discontinued the training sessions after a total of 4 hours training. The subject's bright disposition and clear articulation suggested that she would be an ideal candidate, however the VIM appeared unable to recognise her speech, possibly due to her very high pitch. Subjects 7 and 9 went on to complete 9 hours training each and then used the voice input equipment to enter data to play games and write short stories. The speed and accuracy achieved was recorded for later analysis.

(f) Results

The experimental data was collected during actual classes and exercises and was therefore influenced by a number of extraneous factors including:-

(i) The content of the exercise.

(ii) The subject's state of health and attitude on the day.

(iii) Other activities in the classroom creating background noise and distraction.

Despite these extraneous factors the data provided by the people conducting the exercises tended to be consistent. The data was summarised and averaged as shown in the table in Figure 9.

(g) Interpretation

Six of the nine subjects were able to enter data with greater than 50% accuracy using their voice. Based on the criteria set out in section 2.4.3, however, only three of the nine were able to enter data with an accuracy which would be likely to be of value in classroom situations.

Subjects 5,6 and 9, who achieved high accuracy, have disabilities which do not affect their vocal tract control,

No. Sex Age	Diagnosis	Student's Alternate Method			Voice	
		Characters per Minute	% Correct	Description	Characters per Minute	% Correct
1 M 10yr	cerebral palsy spastic quad.	9	98	by hand typewriter	15	66
2 F 10yr	cerebral palsy spastic quad.	11	95	as above	15	56
3 M 10yr	cerebral palsy spastic quad. -marked ataxia	16	98	as above	20	70
4 F 9yr	cerebral palsy spastic quad.	16	98	head mounted pointer	24	59
5 F 15yr	arthrogryposis multiplex congenita	53	97	mouth held pointer	34	95
6 F 15yr	Incomplete Traumatic quadraplegia at C4	31 38	100 100	by hand keyboard hand writing	12	93
7 M 15yr	cerebral palsy athetoid quad.	20	100	hand held pointer and keyguard	10	30
8 F 7yr	osteogenesis imperfecta	28	100	hand writing	35 12	25 44
9 M 14yr	duchenne muscular dystrophy	32	100	hand writing *	19 9	82 100

FIGURE 9. Results of data entry trials carried out by nine students.

* Student 9 was able to continue handwriting for only a few minutes due to fatigue.
Data for subjects 1 to 5 was derived by requesting subjects to perform 2 minute data entry exercises at maximum speed. The data for subjects 6 to 9 was derived by requesting subjects to enter the alphabet at maximum speed, and/or by observing normal classroom exercises.

however subjects 6, and to a lesser extent 9, do have reduced breath control. Subjects 5, 6 and 9 have other modalities which allow more rapid data entry and a higher accuracy.

Subject 6 disliked voice input due to the frustration involved and expressed a strong preference for her other modalities. Subject 5 expressed some interest in voice input although enthusiasm for the modality varied. It was felt at the time that voice input might be a useful "back up" modality for use when muscular pain prevented the use of a mouth held pointer for keyboard entry. However this proposal was not taken up and subject 5 no longer uses voice input.

Subject 9 is able to write faster and more accurately by hand than by voice, however he is only able to continue writing for a few minutes due to fatigue. Voice input is the only modality available which allows prolonged data entry. The psychologist assessing the subjects performance reported that he enjoyed writing sentences and playing games using the voice input modality, and that voice input "would appear useful for communication purposes".

Hypothesis H3 is therefore found to be partially supported because some members of the trial group are able to enter data into a computer using the VIM.

Hypothesis H4 is not supported since the VIM is not the preferred modality for any of the subjects in classroom environments. The VIM is the preferred modality for one subject for personal activities, however the speed of data entry would be too slow for a classroom environment.

5.4 Phase Three - Speech Therapy

(a) Hypotheses tested:-

H5. That the VIM may reliably discriminate between sounds representing common articulation errors when the VIM compares the speakers sounds with reference patterns produced by someone other than the speaker.

H6. That the VIM may reliably discriminate between sounds representing common articulation errors when the VIM compares the speakers sounds with reference patterns produced by the speaker.

(b) Subjects

Two experiments were used to test hypotheses H5 and H6. The first experiment employed four subjects, the author (H) and three other able bodied adults (M, S and J) aged between 22 and 32 years of age. H and M are males and S and J are females. The second experiment involved five subjects, the author (H) and four other able bodied adults (P, F, L and K) aged between 25 and 32 years of age. H and P are males and F, L and K are females.

English with an Australian accent is the natural language for all the subjects. Apart from this, however, there was no attempt to select subjects with similar, or distinctive voices. All subjects, except the author H who had already extensively used the equipment, were given a short introduction to voice input techniques prior to testing and were co-operative users of the equipment.

(c) Design

A fundamental requirement of the VIM sound recogniser is that it must have an example of the sound pattern in its memory if it is to recognise the sound. Therefore, if the VIM is to be used to discriminate between correct and incorrect pronunciations, it must have an example of each stored in its memory. Generally, there would be an expectation that the VIM's performance would be better when the reference patterns and the incoming patterns for examination are produced by the same person. In the clinical situation this may be difficult to achieve since it may be undesirable, or impossible to have the client consciously repeat the mispronunciation for storage into the VIM. Hypotheses H5 and H6 examine the discrimination performance of the VIM and the effect of the source of the reference patterns.

Two sets of experiments were carried out to quantify the accuracy with which the VIM can discriminate between sounds representing articulation errors. In each experiment, twenty sets of similar sounds (generally varying by one phoneme) were uttered by the subjects and analysed by the VIM. A second person, confirmed that the sound uttered was correctly enunciated and recorded the VIM's decision. The recorded results have been used to test hypotheses H5 and H6. In both experiments two cases were examined.

In case one the VIM discriminated between sounds, when the speaker's sounds were the reference patterns and in case two when some other subject's sounds were the reference patterns.

The difference between the two experiments was in the selection of sounds. Experiment one examined the performance of the VIM when the reference patterns were made up of five different words, four of which were examples of different types of pronunciation errors. Experiment two examined the performance of the VIM when the reference patterns were made up of two different words, so that only one pronunciation error was present in each reference set.

Appendix A6-d lists the sets of sounds used during experiment one. It may be seen that files 0 to 13 are made up of a primary sound and four sounds which are different to the prime sound by one phoneme. The changes in the sounds represent changes which might occur due to mispronunciation. The examples represent errors in production, both place and manner, and in voicing and in final consonant deletion. File

14 has examples of changes in blended sounds. File 15 has examples of changed vowels which were included to test the VIM's sensitivity to vowel changes. Files 16 to 19 contained only two words. Each file was an example of one type of pronunciation error.

Experiment one was set up with this pattern of sounds because it was felt that articulation training programs such as "Voice Tutor" would be most useful if supplied with reference libraries complete with the several errors which might occur.

Appendix A6-e lists the sets of sounds used during experiment two. It may be seen that each file contains only two sounds. Each file represents one of the four types of pronunciation errors considered. The sounds used in experiment two were the pairs of sounds from experiment one where best discrimination performance was achieved.

(d) Apparatus

The VIM was set up in the Apple IIe computer in accordance with the manufacturer's instructions. The table mounted microphone supplied with the VIM was used throughout phase three.

The software called "Voice Tutor" (Appendix A4) was used throughout the phase three experimental work. "Voice Tutor" was used due to its rapid training technique, rapid file manipulation, and because the display shows which sound was actually recognised by the VIM.

(e) Procedure

The procedure adopted for experiments one and two were similar and followed the steps listed below. (See Appendix A4 for detailed description of the activities mentioned).

(i) The subject was introduced to the equipment and trained in correct operational techniques.

(ii) The subject trained the sounds listed in Appendix A6-d or A6-e into the computer as reference patterns. A short test was performed after training in each file to ensure correct training. Rest periods were taken as required to avoid effects due to fatigue.

(iii) Following a suitable rest period after the training session the computer's disc drive was loaded with either the reference patterns generated in step (ii) or by reference patterns generated by one of the other subjects. Whenever possible the source of the reference patterns was withheld from the subject to avoid the subject anticipating a particular result and changing his pronunciation accordingly.

(iv) A file of sounds was then entered into the computers reference area and the subject spoke the words in that particular file. A second person, generally the author, listened to the subject to confirm the pronunciation was correct and to record the computer's response. Appendix A6-f is a typical results sheet. When all sounds on the file were tested, the next file was loaded and tested until all 20 files were tested.

(v) Step (iv) was repeated using a different set of reference patterns. All subjects carried out step (iv) at least twice - Once with their own reference patterns and at least once with a set produced by another subject of their own sex.

(f) Results

A standard test form was completed during the test runs to record the sound uttered by the subject and the sound recognised by the VIM. A typical results sheet is shown in Appendix A6-f. Data from these test sheets has been summarised on the tables in Appendices A6-g and h.

Appendix A6-g shows a summary of the results obtained in experiment one. The table in A6-g shows whether the computer recognised the spoken sound correctly on the first attempt for files 0 to 15. A "1" is shown when the combination of speaker, reference patterns and spoken word led to a correct recognition by the computer.

Appendix A6-1 shows a summary of the results obtained in experiment two. The table in A6-1 shows the number of times that the computer correctly recognised the spoken sound out of three attempts for each combination of word, speaker and reference pattern.

(g) Interpretation

In experiment one of phase three the VIM was used to compare a spoken sound with five reference patterns, four of which represent possible pronunciation errors of the fifth sound. The following points summarise the experimental data in Appendix A6-g:-

(i) There were 175 correct recognitions by the VIM out of 300 attempts (i.e. 58% correct) when the speaker's sound is compared with his own reference patterns.

(ii) There were 133 correct recognitions out of 300 (i.e. 44% correct) when the speaker's sound is compared with another person's reference patterns.

Clearly it would be difficult to use the VIM to supervise articulation training exercises if its discrimination accuracy is similar to the results which would be obtained by flipping a coin. It was felt that performance might improve if the VIM only had one type of error represented in each set of reference patterns. This was tested by experiment two. The sounds used during experiment two were the pairs of sounds identified by experiment one as being most easily recognised by the VIM.

The data obtained during experiment two of phase three is listed in Appendix A6-h. The methods described by Walpole (79) were used to analyse the data in A6-h to produce the table in Figure 10.

Inspection of columns 6 to 11 of the table in Figure 10 reveals that the VIM's average accuracy varied from 7% to 26% correct when the sounds being analysed were compared with reference patterns produced by a different speaker. Consideration of the 95% confidence level yielded a range of accuracy from 0 to 37% correct. One reason for the very low figures is that rejection errors were counted as errors. Even so, it is clear that the VIM cannot reliably discriminate between sounds representing common articulation errors when the VIM compares the speaker's sounds with reference patterns produced by someone other than the speaker. Therefore Hypothesis H5 is not supported.

Columns 1 to 5 of Figure 10 show that the accuracy of correct recognition of the VIM when the speaker's sounds are compared with his own reference patterns varies between mean values of 95% to 56%. However the data also shows that the VIM's performance is dependent on the particular speaker. The 95% confidence limit for the mean values varies between ± 6% and ± 13%. Comparing the subject with the highest correct recognitions with the subject with the lowest correct recognitions reveals with a 95% confidence that the correct recognition rate for the two subjects is different by 39% ± .5%.

COLUMN NUMBER		1	2	3	4	5	6	7	8	9	10	11
SOUND PRODUCED BY:		H	P	F	L	K	H	P	F	L	F	K
REFERENCE PATTERNS PRODUCED BY:		H	P	F	L	K	P	H	L	F	K	F
MEAN SCORE OF CORRECT RECOGNITION OUT OF 3 ATTEMPTS		2.85	2.40	1.67	2.12	2.40	0.62	0.52	0.20	0.77	0.22	0.57
STANDARD DEVIATION		0.58	1.08	1.22	1.16	0.87	0.95	1.04	0.56	1.07	0.70	1.03
95% CONFIDENCE LIMIT	UPPER	3.03	2.73	2.05	2.48	2.67	0.92	0.85	0.37	1.11	0.44	0.90
	LOWER	2.67	2.06	1.30	1.77	2.13	0.33	0.20	0.02	0.44	0.01	0.25
MEAN AND 95% CONFIDENCE LIMITS ON CORRECT RECOGNITION EXPRESSED AS %	UPPER %	101	91	68	83	89	31	28	12	37	15	30
	MEAN %	95	80	56	71	80	21	17	7	26	7	19
	LOWER %	89	69	43	59	71	11	7	1	15	0	8
SUBSTITUTION ERROR RATE %		4	8	4	3	7						

FIGURE 10. This table summarises the experimental data in Appendix A6-h. It illustrates the importance of the origin of the reference patterns to the discrimination accuracy of the VIM.

Appendix A6-e contains the test utterances used to obtain the above data.

The data discussed in the last paragraph refers to correct recognitions where the VIM correctly matches the incoming sound with the correct reference pattern. There are two main types of errors (21a):-

(i) Substitution errors, or the mistaking of one word for another.

(ii) Rejection errors, or refusal to classify a word.

The bottom row of Figure 10 shows the substitution error rate for the five subjects. It may be seen that the rate does not exceed 8% and in three of the five cases does not exceed 4%. It is suggested that with time a user would tend to improve his correct recognition rate by reducing the rejection error rate. This proposition is not tested in the experiment however it is supported by the results obtained by subject H (the author). Subject H achieved a 95% correct recognition rate which is significantly higher (15% to 39%) than the rates achieved by the other subjects however the substitution error rate is approximately the same for all subjects. It is suggested that the author's extensive use of the VIM has allowed him to control his pronunciation to reduce the rejection error rate. In other words he has learnt to pronounce words so that the characteristics examined by the VIM remain more constant. This tendency supports the idea of using the VIM for articulation training.

If it is true that use of the VIM would decrease the rejection error rate so that the correct recognition rate would exceed 90% then an error tolerant articulation program, such as "Voice Tutor" Appendix A4, might be used with positive results.

In analysing the results to test hypothesis H6 the following points may be made:-

(i) The substitution error rate of 4% for subject H combined with a 1% rejection error rate represents reliable recognition.

(ii) The substitution error rates of 8% and 7% for subjects P and K combined with rejection error rates of 12% and 13% represent marginal recognition reliability. However an error tolerant "games" program could be used successfully with these errors rates.

Therefore hypothesis H6 is found to be supported for some people.

6. DISCUSSION

6.1 The Primary Hypothesis

The primary hypothesis considered by this thesis is: "That the VIM sound recogniser is of practical use to members of the community served by Regency Park". The test of the hypothesis has been whether the disadvantages inherent in the employment of the device outweigh the benefits derived by the user. This test has been carried out by performing appropriate steps to implement the VIM and to then monitor the usage. At the present time, after an extensive period of evaluation and implementation, the VIM is not being used on a regular basis, and there are no definite plans to use it. The primary hypothesis, therefore, is not supported.

While the test shows that the VIM is not of practical use at Regency Park at this time, the result should be qualified to some extent. The following sections, 6.2 and 6.3, take up this issue and use the experimental data to develop methods which may assist the therapist in predicting when voice input should be considered.

6.2 Data Input

The experimental work carried out to test the secondary hypotheses H1 to H4 shows clearly that speech may be used to enter data into computers by both the able bodied and the disabled. However, the test of the primary hypothesis has shown that voice input has not been adopted because the disadvantages outweigh the benefits. This conclusion was drawn by the students, teachers and therapists based on a number of different factors which are often difficult to quantify. Sometimes this decision is made subconsciously or by default, for one reason or another there is a general lack of enthusiasm for a technique and it simply falls into disuse. It is important however to understand these factors and develop yardsticks to measure them so that in future modalities may be more closely matched to clients. Initially, below, the author identifies the primary disadvantages of the voice input equipment used in this study. Voice input is then discussed in terms of the characteristics developed in section 2.4.2.

All students reported varying degrees of frustration due to the limited accuracy of the equipment. This problem was displayed in two ways:- (a) a percentage of errors and (b) the tendency to repeatedly ignore or produce an error for one or two particular sounds during a session. The second of these proved to be most frustrating since the user would begin to expect an error when saying the word, and would then have great difficulty in correcting the entry. This frustration was most apparent for subjects five and six who were using word processing programs and were accustomed to modalities which allowed them certain access to the whole keyboard. The accuracy problem was generally the main problem, however the other aspects listed below also reduced the desirability of the modality.

In at least two cases the speed of data entry with voice input was significantly inferior to the subject's normal modality.

In one case, subject 5, voice input appeared to be desirable as a back up modality since the subject's normal modality sometimes causes muscular pain. However, the high cost of the equipment, and the lack of portability, and therefore utility, made the option impractical.

Problems due to fatigue and environment were generally of little concern to the subjects compared with the other problems already discussed. Subjects six and nine were exceptions however. Subject six reported a high level of fatigue, although this may well have been related to her frustration due to the accuracy problem. Subject nine, on the other hand, reported significantly reduced fatigue when using voice input when compared to his normal modality.

While these problems are significant they must be seen relative to the advantages and disadvantages of the subjects other modalities. For example, subject 9 encountered similar problems to some of the other subjects however he preferred, and enjoyed, speech input for prolonged data entry. This was because his normal modality became impractical due to fatigue. The following compares voice input with the subjects other modalities, using the characteristics defined in section 2.4.2. Figure 8 presents a summary of the VIM's performance as a data input modality for the disabled. The data could be used to assess the suitability of voice input for a client.

(a) Accuracy

(i) Presentation

Seven of the nine subjects tested had access to keyboards using their alternate modality. It seemed likely that the other two subjects could manipulate a keyboard if required. Therefore there is no significant difference in this aspect.

(ii) Rates of correct data entry

In all cases voice input produced a lower rate of correct data entry than the subject's other modality. The correct data entry rate did not fall below 95% for any of the subjects using their normal modality. However the correct data entry rate was below 80% for all subjects with impaired speech.

(b) Power

Figure 11 below compares the power of the various modalities used during the experimental phases. The vocabulary size for the VIM is 80 and a vocabulary size of 100 is taken for standard keyboard entry and for handwriting. It should be stressed that the equation for "Power" was derived arbitrarily in section 2.4.2, and is included here to illustrate a proposed methodology for comparison of different modalities.

Subject Number	Voice	Normal modality
1	0.35	0.30
2	0.35	0.33
3	0.40	0.40
4	0.44	0.40
5	0.52	0.73
6	0.31	0.62
7	0.28	0.45
8	0.31	0.53
9	0.39	0.57*

* Note: The normal modality power for subject number 9 may only be maintained for a few minutes. It then drops to zero due to fatigue.

Figure 11. Comparison of modality power. Data taken from figure 9.

Figure 11 shows that the power of the subjects normal modality is significantly greater than for voice for subjects five to nine. The modality power is approximately the same for subjects one to four, whose hand movement is affected by their disability.

Subject nine's normal modality has significantly higher power until fatigue prevents further use (after a few minutes). Voice input would be appropriate here, however the benefits which might be gained from the irregular times when protracted data entry is required have not justified the regular use of the equipment.

(c) Convenience

(i) Cost

The VIM cost approximately $1700 and requires a computer and printer costing, say, $2500 to operate and print out information. Although some subjects used pencils and inexpensive typewriters, to fairly compare the presentation and utility of the modalities it is assumed that in the longer term the subjects would operate a word processor residing in a personal computer. On this basis voice input costs approximately $1700 more than the other modalities considered.

(ii) Portability

The VIM resides in an Apple IIe computer which may be moved, but is not readily portable and must be connected to a mains power supply. This, however, would be typical of a personal computer offering word processor functions.

(iii) Frustration

Several students reported frustration when trying to enter a particular sound into the VIM which the VIM would not accept. They do not suffer the same frustration using their other modalities since all characters are accessible.

(iv) Environment

The speech input device would require a relatively quiet environment during operation. In public places the user might find operation of the device embarrassing. The user's speech might also be intrusive and distracting in, for example, a classroom environment.

(v) Fatigue

Generally, but not in all cases, the students involved in the testing found voice input to require less effort than their normal modality. This was particularly important for subject nine because his normal modality cannot be used for more than a few minutes. Effort required was also an important aspect for subject 5 who sometimes has difficulty using her normal modality because of muscle pain.

1. ACCURACY

(a) Presentation
The VIM enters data to a computer so that all physical output is typed.

(b) Typical rates of correct entry (from experimental data) are shown in the table below.

EXTENT OF SPEECH IMPAIRMENT	% OF ENTRIES RESULTING IN A CORRECT ENTRY	POWER
NO IMPAIRMENT	80% - 100% *	0.31 - 0.52
MODERATE IMPAIRMENT	50% - 80%	0.35 - 0.44
SEVERE IMPAIRMENT	0% - 50%	0 - 0.28

* NOTE: One subject with no speech impairment was only able to achieve a 40% correct entry rate for no apparent reason.

2. POWER

Vocabulary size = 80. Rates of entry from experimental data. The power is shown in the table above. Note that the values for power are obtained from an unproved formula.

3. CONVENIENCE

(a) Cost. Approx $1700 for the VIM and $2500 for the host computer.

(b) Portability. Not portable, mounted in a personal computer.

(c) Frustration. Several subjects reported frustration when unable to enter particular characters.

(d) Environment. A relatively quiet environment is required for operation.

(e) Fatigue. Several subjects reported reduced effort required to enter data.

FIGURE 12. Data Input Performance Summary For The VIM Voice Input Device.

6.3 Speech Therapy

The application of voice input equipment for articulation training is at a very early stage of development. Several areas would require investigation before forming a conclusion on whether voice input equipment would be suitable for the application. The following three areas are identified:-

(a) The equipment should be examined to find whether, and in what circumstances, it can discriminate between correct and incorrect articulation. Issues which should be included are:-

  (i) The accuracy of the equipment, what percentage of times does the equipment correctly assess the sound made.

  (ii) The type of error made, whether the sound is wrongly assessed (i.e. substitution error) or the sound is ignored (i.e. rejection error). The ratio of substitution errors to rejection errors would be important when considering the appropriate feedback to the client.

  (iii) Whether the reference patterns would have to be produced by the client. It would be desirable to have the therapist generate the reference patterns.

(b) It would be important to find, and define types of people who would operate the equipment to produce acceptable accuracies. Considerations would include:-

  (i) The quality of the person's speech and the type of impairment. For example, one of the students involved in the data entry testing spoke with clear and easily understood speech, but was unable to obtain the expected good performance from the equipment. It was felt that this may have been due to the student's high pitched voice. Other researchers (23) have reported similar problems. The equipment also tended to produce poor results for students with poor breath control. This was probably due to random breathing noises and inconsistent length and loudness of speech.

(ii) The cognitive ability and maturity of the individual. One of the things that became clear from testing the equipment was that it is very sensitive to changes in pronunciation due to the emotional state of the operator. Individuals who allow frustration to affect their speech following a misrecognition tend to suffer a decrease in performance. This increases their frustration and further decreases their performance. Most test subjects were able to understand the problem and many adopted a bland, consistent pronunciation, to overcome the problem.

(c) Although repetition drills are often employed, their value for improvement of articulation is questionable for many clients depending on their diagnosis and personality. Voice recognition equipment may offer the potential for a great increase in the number of drills performed, however the potential of this for improvement of articulation is not known.

The experimental work in this thesis has concentrated on the ability of the equipment to discriminate between correct and incorrect sounds as described in (a) above.

The interpretation of the results in chapter five discusses the results at length and suggests that the VIM might produce acceptable levels of discrimination for articulation training in some cases. The client would, however, be required to enter examples of the correct and incorrect pronunciation into the VIM prior to the articulation training session. A tape recorder might assist in this process, however this has not been tested.

Experimental work has only been carried out with able bodied adults simulating pronunciation errors. The results suggest that the VIM would classify sounds in three ways:- 1. Correctly assess the sound, 2. Incorrectly assess the sound or 3. Find that the sound does not match any sound in memory (rejection error). Although the percentage of correct assessments varies greatly from person to person it appears that incorrect assessments form a small relatively fixed proportion. The variation is made up by the rejection error. So in practice the computer would tend to ignore the client until the speech pattern entered matches a stored pattern. It would then be likely to correctly assess the sound. It is suggested that the client would tend to learn the sound required by the computer and modify his speech accordingly.

Although the results were encouraging it was felt that it would be inappropriate to trial the equipment with children because they might blame themselves for machine errors leading to a decrease in self esteem.

7. CONCLUSION

The following conclusions are drawn from the preceding chapters.

Data Entry

The voice input equipment examined is generally not of practical use for data entry at Regency Park. This was shown by the fact that the users of the equipment stopped using the equipment. The equipment was technically supported and enthusiastically introduced in appropriate situations. It is concluded that the reason for the lack of continued use is that the effort required to operate the equipment does not justify the benefits gained. This is an entirely subjective decision on the part of the user. The fact that so many users reached the same decision is significant. The equipment failed to achieve a satisfactory accuracy when used by able bodied adults simulating dysarthric speech, and when used by students with impaired speech.

In spite of the negative result, voice input of data should not be neglected. The literature search revealed cases where voice input is of practical and significant benefit. The unique nature of the abilities and personalities of disabled people leads to the situation where a modality such as voice input is suitable for a very small number of people. Data gathered during the experimental work is presented to allow a therapist to objectively decide whether voice input is a desirable option for a particular person. An unproved methodology for comparison of different modalities is proposed.

The literature search also reported significant recent advances in voice input technology. These advances will make voice input of data a more attractive option for both the disabled, and able bodied communities as time goes by.

Speech Therapy

The results obtained from experiments concerning articulation training using voice input equipment varied greatly from person to person. However it is concluded from the results that the equipment would not discriminate with sufficient accuracy when the client's voice is compared with voice patterns generated by another person.

Insufficient experimentation has been carried out to determine whether the equipment will offer adequate accuracy when comparing the client's voice with his own reference patterns. Although the error rate varied markedly between subjects, the substitution error rate was small and consistent with the variation being made up of rejection errors. The subject with the greatest experience of use of the equipment produced significantly reduced rejection errors. This suggests that articulation consistency may be a learned ability which may be developed using the equipment.

It is therefore concluded that the results justify further experimental work in this area. It is suggested that future work would initially confirm that subjects rejection error rate improves with use of the equipment. The substitution error rate would then be considered to determine whether an interactive game could be successfully played to motivate the user.

APPENDIX A1

IMPLEMENTATION OF A VOICE ACTIVATED COMPUTER SYSTEM FOR TEACHING AND COMMUNCIATION.

Holmes, W., Andrews, P., Seeger, B., Radcliffe, D.

A paper presented at TADSEM '85 - Australian Seminar on Devices for Expressive Communication and Environmental Control. Paper published by Technical Aid to the Disabled, Box 108, Ryde, N.S.W. 2112

IMPLEMENTATION OF A VOICE ACTIVATED COMPUTER SYSTEM FOR TEACHING AND COMMUNICATION

Mr. W Holmes*, Mr. P Andrews**, Dr. B Seeger**, Dr. D Radcliffe*

** Regency Park Centre for Young Disabled
* University of Adelaide, Department of Mechanical Engineering

ABSTRACT

Several relatively inexpensive speech recognition systems, suitable as input systems for personal computers, are now commercially available. One, the Voice Input Module (VIM, from Voice Machine Communications) was selected for development for use with disabled young people at the Regency Park Centre for Young Disabled. The VIM consists of a microphone, a plug-in board to fit an Apple II or IIe computer and software to input, edit, test, and use an operator's vocabulary as an input signal to a computer.

An overview of the hardware and software of the VIM is presented. The VIM software is designed for able bodied computer hobbyists. Modifications to the basic software to make the system user friendly so that people with disabilities can have voice only input (ie. no input whatsoever from the keyboard) have been developed by us and are outlined in the paper. The VIM may be trained to recognise any type of utterance. Preliminary tests showed that the International Civil Aviation Organisation (ICAO) alphabet (alpha, bravo, charlie) had significantly better recognition rates than the conventional a,b,c.

There are a number of children at the centre who could benefit from a voice activated assistive communication device. A trial group of five children has been selected to provide preliminary evaluation of the potential of VIM. Existing educational software (eg. Mathvader) has been used with the VIM to provide appropriate programs for their school level. Our initial experience, both successes and problems, are presented.

INTRODUCTION

Several hundred automatic speech recognisers have been developed since 1952 when Bell Telephone Laboratories developed the first speech recogniser [1]. Many have been successfully used in factories for stock inventory and data entry where more traditional methods have been inappropriate. Recent advances in large scale integrated circuits and the mass production of home computers has radically reduced the cost, and increased the availability of these devices. Several authors have suggested that voice input has potential for the handicapped to improve the quality of life [2]. An impressive example is provided by Bowe [3] who reports that a computer programmer who recently became a quadraplegic has been able to continue his career by voice control of a computer terminal at home. However Poulton [4] cautions that present recognisers are limited in scope and may only be suitable where demanded by the constraints of work or the classroom.

A recent study at the Regency Park Centre for Young Disabled [5] identified a group of twelve children with poor handwriting skills but good voice control. This paper is a progress report on current work aimed at introducing voice input to this group. Topics covered include equipment assessment, initial software modifications, application programs and preliminary results of equipment trials with five students. It is concluded that voice input will be a valuable technique for accessing computer programmes for a limited population of disabled people.

VOICE INPUT MODULE - VIM

In 1984 Regency Park purchased a voice input module (VIM) from Voice Machine Communications [6]. The VIM is a speaker dependent, isolated word recogniser which acts as a keyboard emulator for its host Apple IIe computer.

The VIM recognises utterances by comparing incoming audio signals with a previously entered set of voice reference patterns. 80 utterances may be recognised with a length of from 0.1 to 1.25 seconds separated by pauses of at least 0.16 seconds. A predetermined set of keystrokes is sent to the Apple II's central processing unit following recognition of an utterance.

Software supplied with the VIM allows the operator to define and manipulate three files. The first file contains a list of user defined words which prompt the user when entering voice reference patterns. The second file contains a list of user defined keystrokes which are sent to the computer when an utterance is recognised. The user speaks into a microphone in response to the prompt words to generate the voice reference patterns which form the third file. The voice reference patterns may be tested for separability and retrained if necessary, and then stored on floppy disk for later use. The operator may also modify the "reject threshold" to define how well an incoming signal must match a voice reference pattern to cause recognition. For example, a reject threshold of 1 will cause the VIM to accept almost any sound and match it to the closest voice reference pattern. If the reject threshold is 127 then the VIM will only recognise sounds identical to the voice reference pattern.

PERFORMANCE TESTING

A performance test was conducted to determine whether the equipment offered sufficient potential to justify further work.

The test was designed to show how often the equipment would correctly identify an utterance, how often an utterance must be repeated to produce recognition and how these are affected by the reject threshold. The test was conducted with one subject familiar with the equipment with a normal adult male voice. The author was required to repeat three sets of words and sounds. These sets were selected to give an intuitive impression of how well the equipment can discriminate between sounds and how effective utterances other than words might prove. The three sets of utterances were;

(1) The alphabet pronounced "ay", "bee", "see"
(2) The ICAO aviation alphabet pronounced "alpha", "bravo", "charlie"
(3) The Morse code pronounced "DiDah", "DahDiDiDi", "DahDiDahDi"

The aviation alphabet enabled the equipment to correctly recognise 95% of utterances. However the recognition rate dropped to 82% for the standard alphabet and 63% for the Morse code style. The above rates were obtained using a reject threshold of 105 which virtually eliminates the need to repeat utterances.

MODIFICATION OF SUPPLIED SOFTWARE

Software supplied with the equipment allows the user to enter new voice reference patterns, vocabulary files, and activate application programmes (ie. any commercially available programme). These activities require physical contact with the computer keyboard and a good knowledge of the supplied VIM software. The software is well documented and easy to use, however it is pitched towards the computer enthusiast rather than a casual user.

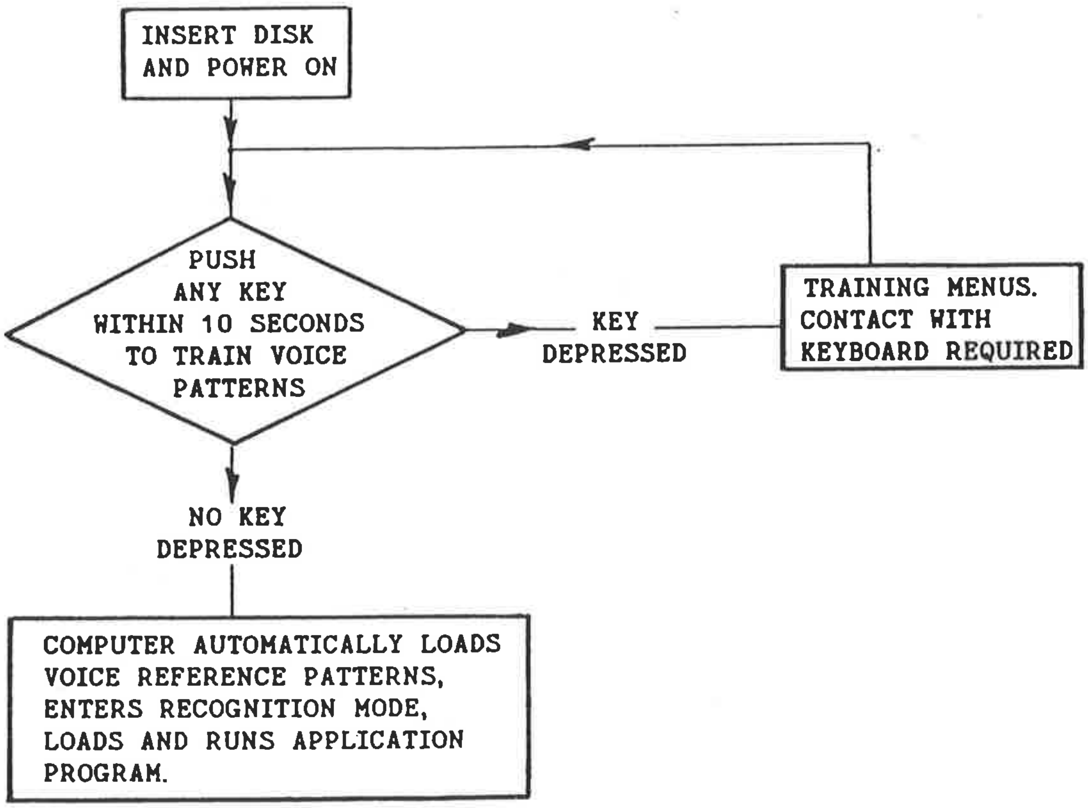
The following strategy and software modifications were adopted to make equipment more compatible with the classroom environment.

(1) Each application program used by a given student is stored on a separate floppy disk which also stores the student's voice reference patterns.

(2) At the beginning of a typical lesson the teacher inserts the student's floppy disk and turns the power on to the computer. The computer automatically loads the voice reference patterns, transfers to recognition mode and runs the application program without further intervention by the teacher. The student can then start using the program with his voice.

(3) Contact with the keyboard is required when first entering the student's voice reference patterns or to retrain single words if required. A prompt displayed during the loading process described above invites the teacher to push any key within ten seconds to go to the training mode. If no key is pushed then the disk loads as described above. If a key is pushed then a menu is displayed which prompts the teacher through the essential training options. Following training the program returns to the load sequence.

The process described above is intended to make the system as simple as possible to operate and is illustrated in the flowchart below.



APPLICATION PROGRAMS

The application program is the software which performs the task required by the user. Word processing programs, arithmetic practice programs and games programs are application programs.

The following five programs have been used extensively during the study and require no physical contact with the keyboard after the initial training is complete.

The programs "Smile" and "Wuzzle" invite the user to count from one to nine objects displayed on the screen. A correct answer is rewarded by a smiling face or a short cartoon. The user is encouraged to try again if the wrong answer is given.

"Spaceship" provides the user with practice in addition of single digit numbers. The program rewards correct answers with a cartoon and invites the user to try again following an error. The correct solution is displayed following a further incorrect answer.

"Mathvader" is a more advanced arithmetic practice program allowing practice in addition, multiplication, division and subtraction of any one or two digit numbers. The user has an adjustable time period to complete calculations and win the game. The game is modelled on the popular arcade game "Space Invaders" and is quite appealing.

"Bank Street Writer" is the word processing program used during the study since it is simple to use and all functions and keystrokes may be used by voice alone. "Xardax" and "Wordstar" were rejected because they require a significantly larger number of discrete utterances to access all keystrokes (upper and lower case) and functions.

EQUIPMENT TRIAL WITH STUDENTS

The equipment trial was designed to assess whether the selected students could use the equipment to provide advantages over other techniques of keyboard manipulation at their disposal. The trial has been conducted in the classroom environment with application programs complimenting the students other academic activities. Generally, the classroom sessions have been coordinated and conducted by one of the special education teachers at Regency Park. The following is a summary of progress to date of five students.

Students One To Four

Students one to four are aged from 9 to 10 years old. Two of the students can count to ten confidently and identify letters of the alphabet. The other two students can add numbers and spell three letter words by sounding the letters. All four students have easily recognisable voices but have difficulty manipulating a keyboard due to spasticity and/or weakness.

The main obstacle to voice input for these four students is that they tend to allow their pronunciation to vary with mood changes. For example, they will be bored during the training session, enthusiastic at the first attempt of an exercise, disappointed if the VIM does not recognise their voice, and frustrated or angry after the next failed attempt. Each of these mood changes significantly changes their pronunciation to the point where the VIM fails to recognise their words. Several therapists have suggested that the mood changes described are typical for children with cerebral palsy.

The student's exercises were timed after they had used the voice input for a total period of about ten hours to derive the results in table one. The table shows that the four students are able to enter data at a faster rate with voice input, however there is an unacceptable level of errors. We feel that this high error rate is due to the student's age and lack of maturity since the quality and strength of their voices appears normal.

Student Five

Student five has good control over her clear voice and is able to achieve good data entry rates with a low number of errors. The table shows that she is significantly faster using a mouth held pointer however she is not able to maintain this high rate due to fatigue and occasional discomfort. The student prefers voice input and is able to use her voice to control and operate a word processing program. We expect that voice input will be a valuable alternative for this student and that she will combine voice and physical contact to obtain her best overall performance.

no. name age	diagnosis	student's alternate method			voice	
		characters per minute	percentage correct	description	characters per minute	percentage correct
1 BN 10yr	cerebral palsy spastic quad.	9	98	by hand typewriter	15	66
2 KH 10yr	C.P.S.Q.	11	95	as above	15	56
3 DF 10yr	C.P.S.Q. with marked ataxia	16	98	as above	20	70
4 GS 9yr	C.P.S.Q.	16	98	head mounted pointer	24	59
5 MR 15yr	arthrogryposis multiplex congenita	53	97	mouth held pointer	34	95

Table 1. Results of trials with the first five students.

The table above was derived by encouraging the students to enter data at their maximum possible rate for two minutes. Other factors which must be considered when assessing the suitability of voice input would include;

(1) Effort and possible discomfort involved.

(2) Long term effects. For example the possible damage to neck muscles and joints due to head mounted pointers or switches.

(3) Workplace ergonomics. For example a student using a mouth held pointer to transcribe material must make gross head movements to read the document and then type the material.

(4) Speed and accuracy over longer periods using voice and their best alternative method.

(5) The student's need for, and capacity to use a non portable computer or typewriter.

FUTURE EQUIPMENT TRIALS

Students one to four appear to be too young to effectively use the equipment due to their high error rate. However their rate of data entry compares well with their best alternate method so we feel that a further trial will be justified in two to three years.

Student five will continue to gain confidence and skill with the equipment and will be encouraged to use strategies combining voice and physical contact.

Five new students have been selected and will work with the voice input equipment during the current school term. Ultimately all students served by Regency Park who may benefit from voice input will be given the chance to try the equipment. We expect that we will find a small group of children who will gain significant benefits from voice input. The software will then be tailored to suit the needs of this group.

CONCLUSIONS

Voice input equipment has been adapted for use by disabled children in the classroom environment, and a trial is underway at the Regency Park Centre for Young Disabled.

Voice input is by no means a panacea as it requires some skill and effort to produce a relatively slow stream of information. However, it will provide a viable alternative for a number of people who are currently unable to operate keyboards to make full use of available technology. As a guide, therapists might consider voice input for subjects with controlled, consistent speech, who have a need to enter data to a non portable computer and currently convey less than thirty characters per minute.

The voice input module costs approximately A$1600 and may be obtained from Voice Machine Communications Inc [6]. We have not tested other brands of voice recognition equipment and therefore cannot recommend this brand over others. However it should be noted that the VIM is supported by utility programs written in BASIC which are readily modified to make the equipment more suitable for particular users.

Software development and equipment trials are expected to continue through to the middle of 1986 when a fully documented software package suitable for use by people with disabilities will be available from Regency Park. The authors welcome any enquires on the subjects discussed above [7].

ACKNOWLEDGMENT

We wish to acknowledge and thank the Apple Education Foundation which has provided funds for the computer equipment used with the VIM.

REFERENCES

1. Lea, W.A. (Editor) Trends in Speech Recognition. Prentice Hall, Englewood Cliffs 1980

2. Damper, R.I. "Voice-input Aids for the Physically Disabled". Int. J. Man-Machine Studies (1984) 21, 541-553.

3. Bowe, F.G. Personal Computers and Special Needs. Sybex Computer Books.

4. Poulton, A.S. Microcomputer Speech Synthesis and Recognition. Sigma Technical Press. (1983)

5. Ahrens, B.A. et al "Assisted Communication Needs of Physically Disabled Children" internal paper at Regency Park.

6. Voice Machine Communications Inc. 1000 South Grande Ave. Santa Ana CA 92705 U.S.A.

7. To contact the authors;
W Holmes. P.O. Box 634 Adelaide 5001
P Andrews, or Dr. B Seeger. The Regency Park Centre for Young Disabled, P.O. Box 209 KILKENNY, S.A. 5009
Dr. D Radcliffe. University of Adelaide, Mechanical Engineering Department. S.A. 5000

APPENDIX A2

TEACHER'S VIM

USER'S

MANUAL.

TEACHER'S VIM

USER'S

MANUAL

System Requirements

Apple IIe with 64k RAM
at least one disk drive
and a Voice Input Module fitted in Slot 4

Copyright

Acknowledgements

VIM is a registered trademark of Voice Machine Communications Inc. Bank Street Writer is a registered trademark of Broderbund Inc.

CONTENTS

Chapter 1 Introduction 01

1.1 VIM and Teacher's VIM 01

Chapter 2 Using Teacher's VIM 03

2.1 Loading and running Teacher's VIM 03

2.2 Training 05
 2.2.1 Single session training 06
 2.2.1 Training one word at a time 12

2.3 Running the application program 13
 2.3.1 Bank Street Writer 13
 2.3.2 The Extended Number Set 13

2.4 Use in the classroom 14

Chapter 3 Advanced Topics 16

3.1 Steps 7 to 10 of the "Train-Test" Program menu 17

3.2 Vocab and voice pattern files 18

3.3 Hands off operation of software 19

Appendix

A.1 Numbered vocabulary lists for Teacher's VIM
 A.1.1 Bank Street Writer 20
 A.1.2 The Extended Number Set 21
 A.1.3 Pania 22
 A.1.4 Star Blazer 23

A.2 Application program. Instructions supplied by Teachers VIM
 A.2.1 Bank Street Writer 24
 A.2.2 The Extended Number Set 25
 A.2.3 Pania 26
 A.2.4 Star Blazer 27

1. INTRODUCTION

1.1 VIM and Teacher's VIM

The Voice Input Module (VIM), supplied by G.C.D. International, allows the user to perform keyboard functions using his voice. The VIM is a speaker dependent, isolated word recogniser. It will recognise up to 80 different isolated sounds spoken by a person who has trained the machine to recognise the particular sounds. When the VIM recognises a sound it enters a predetermined set of keystrokes into the Apple computer.

The VIM is supported by an extensive utility program which is used to train the VIM with the user's voice patterns. The user may select options through various menus to tailor the vocabulary in the VIM to their particular needs. We have found that in the early stages, the flexibility of the supplied utility program is generally not required and may add to the workload of the teacher or therapist. The manual supplied with the VIM should be consulted for a full description of the VIM's options and for instructions on installation of the hardware.

Teacher's VIM is a modified version of the VIM's Voice Utility Program which reduces the number of options, eliminates the need for file manipulation and prompts the user through the required steps. This reduces the time required to learn to use the VIM and assists the teacher in applying the VIM in the classroom. This manual supplies sufficient information to use Teacher's VIM, however, the user is invited to read pages I1, I2, I3 and IV1 and IV2 of the VIM manual for a more complete understanding of the VIM and for some suggestions on the use of the microphone.

Sections 1 and 2 of this manual provide sufficient information to allow regular classroom use of the equipment. Sections 1 and 2 assume little knowledge in the use of computers, however section 3 assumes a good knowledge of the VIM manual and the Applesoft BASIC language. Section 3 will not be needed until the student has reached an advanced stage with Teacher's VIM.

2. USING TEACHER'S VIM

2.1 Loading and Running Teacher's VIM

The VIM hardware should be installed in the Apple computer in strict accordance with the instructions supplied in the VIM manual. Set up the computer and disk drive(s) and monitor as shown in the manuals supplied with the equipment.

The initial part of the Teacher's VIM program loads information off the disk and asks the user whether training of voice patterns is required. If training is required then a key is depressed and training begins. If training is not required then prompts are shown which assist the user in running the application program.

The following are detailed instructions for loading and running Teacher's VIM. Insert one of the Teacher's VIM disks into disk drive number 1. Turn the power to the computer on (or press "Control-Open Apple - Reset" simultaneously if the power is already on). The computer will start loading information off the disk. After 40 seconds, the display shown in Figure 2 will appear :

Figure 2. Prompt

HIT ANY KEY WITHIN 10 SECONDS TO
TRAIN THE DISK WITH YOUR VOICE

If the user wishes to train voice patterns into the computer (this will be the case if you have not used the disk before) then any key should be typed within 10 seconds. If a key is typed then the "Train-Test Program" menu (Figure 3) will be displayed (see Section 2.2 for detailed instructions in the use of the "Train-Test Program" menu). If no key is pressed during the 10 seconds then the computer assumes that the voice patterns already loaded are correct. In this case, the computer displays a detailed set of instructions on the running of the application program (see Section 2.3).

2.2 Training

The "Train-Test Program" menu may be used in two ways. All the sounds may be trained during a single session, or the sounds may be trained one at a time. If you elect to train all the sounds in one session, but fail to complete the exercise, then the sounds already trained will be lost. A person capable of uttering words consistently every few seconds for about fifteen minutes should be able to train all the sounds in one session. An adult, able bodied person should have no difficulty with this exercise.

The following two sections work through these two methods of using the "Train-Test Program" menu.

Figure 3. The "Train-Test Program" Menu

```
            ******************
            TRAIN-TEST PROGRAM
            ******************

       INITIAL TRAINING PERFORM STEPS 1 TO 6
            -  -------------------------

            1 - NEW TRAINING
            2 - ADDITIONAL TRAINING (TWICE)
            3 - TEST SEPERABILITY
            4 - RETRAIN SINGLE WORDS LISTED
            IN 3. REPEAT 3 AND 4 UNTIL ALL
            WORDS ARE SEPERABLE. THEN GO TO 5
            5 - SAVE VOICE PATTERNS ON DISC
            6 - RUN APPLICATION PROGRAMME
            7 - READ REJECT THRESHOLD
            8 - SET REJECT THRESHOLD
            9 - TEST RECOGNITION
           10 - AVIM UTILITY MENUS
```

2.2.1 Single Session Training

This section describes the use of the "Train-Test Program" menu to train all the sounds on the Teacher's VIM disk in one session.

As described in detail in Section 2.1 : insert the Teacher's VIM disk, turn the power on, and press any key when prompted. This generates the display shown in Figure 3. The step numbers in the following refer to the numbers on the left side of the menu in Figure 3.

STEP 1

Type 1 on the keyboard and press return. A warning that the current voice patterns will be lost will be displayed. Press any key except "return". "Return" will simply return you to the "Train-Test Program" menu. Another display will state that a new training is under way, press any key except "return" to continue. The display in Figure 4 will appear.

Figure 4. Training Prompt

PLEASE SAY

CORRECTION

Say "correction" into the microphone. If the display remains the same and does not blink, then say "correction" again. Check that your lips are 25mm from the microphone and use a clear conversational voice. When the VIM receives a distinct sound it will display the next word to be trained. If the VIM receives two distinct sounds, for example "correc" "tion", then it will associate "tion" with the second prompted word. The third prompt would then be displayed. Don't worry if this happens because the program makes three training passes of the vocabulary. Errors made during the first pass will be corrected later.

After all the words on the vocabulary are trained for the first time, then the program automatically returns to the beginning and prompts the word "correction". The rest of the vocabulary is then prompted in the same sequence. The VIM compares the second pronunciation with the first. If the pronunciation varies by a significant amount then the VIM prompts a request to say the word again. It will continue to do this until a consistent sound is made.

Beginners often say the words in the first training pass with a stilted pronunciation and then relax during the second pass, so that a different sound is made. It is difficult to pronounce sounds consistently if they are spoken with a forced or unusual emphasis. This is why we recommend that you use your normal conversational voice.

Following the second training pass, the program returns to the word "correction" and makes a third training pass. After the third training pass, the "Train-Test Program" menu is displayed. If you wish to discontinue the session at this stage, then type 5 "return" (see STEP 5 for details) which saves your voice patterns on the disk. After the disk drive stops, you may turn off the computer and remove your disk.

STEP 2

If the computer is off then go through the procedure described in 2.1 to obtain the "Train-Test Program" menu. Your voice patterns trained during Step 1 will have been automatically loaded into the computer during the start up.

Although not essential, best results are obtained by training the vocabulary at least five times. Three training passes have been made during Step 1. To perform additional passes, type **2** and "return". A display stating "Additional Training" will be shown. Press any key except "return" to continue. Press "return" if you have changed your mind, this will return you to the "Train-Test Program" menu. A single training pass, similar to the passes made in Step 1 will take place. Following the training pass, the "Train-Test Program" menu will again be displayed. Type **2** "return" again to perform the fifth and final training pass. When the "Train-Test Program" menu is showing, you may press **5** "return" (see STEP 5 for details) to save the voice patterns on the disk and turn the computer off if you wish.

STEP 3

With the "Train-Test Program" menu displayed, type **3** "return". A display will request you to type the prompt device number. Type **1** "return". The computer will automatically compare all the sound patterns that are stored in memory. A display will show the prompt words associated with any sounds that the computer has found are similar. On some Teacher's VIM disks, the same prompt is repeated several times so that naturally, the spoken sound will be similar. This is normal and will not cause a problem. However, if the computer associates different prompts with the same sound, then the computer is likely to give errors.

Inspect the generated list. Each line of the list will contain a pair of prompt words. Ignore pairs which are made up of the same prompt word. Take a note of the number and spelling of the prompts in the other pairs, these words should be retrained since errors would result from their use. Press any key after taking a note of the prompts as described above. The "Train-Test Program" menu will now be displayed. If some word must be retrained, then go to Step 4. If no words need to be retrained, then go to Step 5.

The above covers all that we need to know about the separability test. Refer to the VIM manual for a deeper understanding of the test and the various options if you wish. The example below (Figure 4) illustrates the use of the separability test for the Teacher's VIM disk "Extended Number Set".

Figure 4. Separability Test - Example

Separability Test

#	REFERENCE WORD	NEAREST NEIGHBOUR	DIFFERENCE
13	Yes	Yes	5
37	Four	Four	5
50	Two	Three	4
79	Return	Return	5

End of Test

In the above example, ignore the first, second and fourth pairs, since the pairs are made up of the same prompt word. Record the number 50 associated with the third pair and record the two words "Two" and "Three". This test has indicated that either the word "Two" or the word "Three" has not been pronounced distinctly enough. Go to Step 4 to correct this problem.

STEP 4

This step is used to train or retrain single words. With the "Train-Test Program" menu displayed, type **4** "return". A display titled "Retraining Mode" will be displayed. Four options are presented. Normally the second option is preferred, so type **2** "return". The other options are discussed in the VIM manual, but are generally not required.

When **2** "return" has been typed, a prompt will appear requesting the word number. Type the number of the word you wish to retrain. If you have just completed Step 3, then you should type one of the numbers you recorded. Refer also to the Appendix Section (4.1) for the numbered listings of the vocabularies on the Teacher's VIM disk. Type "return" following the number. The display will prompt you to say a word. If your pronunciation is clear and consistent, then the word will be prompted three times. If the pronunciation is not consistent, then the word will be repeated several times. After the word has been trained, the prompt requesting the next word number will reappear. If you have another word to train, then enter the number of the word. If your training is complete, then push "return" which will return you to the "Train-Test Program" menu.

Now repeat Step 3 to ensure all words have been trained correctly.

If the above use of Steps 3 and 4 fail to produce separable sounds, then consider the following two options :

Option 1 : In Step 3, we recorded the number of one of the two prompt words in the error pair. However, the pronunciation of the word we recorded (the first word of the pair) may be correct. The pronunciation of the second word in the pair may be incorrect. In the example in Step 3, we recorded 50 which assumes the pronunciation of the word "Two" is incorrect. However, the pronunciation of "Three may have been incorrect. In this case, refer to Chapter 4 to find the number of the word "Three".

Option 2 : The words may be unsuitable for the particular person. For example, the word "Alpha" may be unfamiliar to the student, leading to inconsistent pronunciation. In this case, choose a more appropriate word and say it instead of "Alpha". If this does not solve the problem, then refer to Section 3.2 of the Advanced Topics chapter for assistance.

STEP 5

The voice patterns trained during the use of Steps 1, 2 and 4 are stored in the volatile memory of the VIM hardware. If the computer is turned off, then the information will be lost. To overcome this problem, Step 5 is used to store the information on the disk.

Before storing information on the disk, ensure that :

a) there is no read/write protect sticker on the disk

b) the Teacher's VIM disk is in the drive

c) the disk drive door is closed

To store the voice patterns on the disk, type **5** "return" when the "Train-Test Program" menu is displayed.

The computer will occasionally detect a problem when writing to the disk and will display a prompt to alert the user. Generally, the user will be prompted to type any key to get to the main menu. The program will return the user to the "Train Vocabulary" menu, which is part of the VIM Utility Program. If this happens, type **4** "return" to save the voice patterns on disk. The next prompt will request you to type the file name on the disk. Type **VOICE PATTERN** "return". Then type **1** "return" in response to the next prompt for disk drive number. Your voice patterns will then load onto the disk. If you wish to return to the "Train-Test Program" menu, then perform the procedure described in section 2.1.

STEP 6

If you have performed the preceding steps then you are ready to run the application program. Steps 7, 8, 9 and 10 are not normally used and are left to Section 3.1 for discussion.

To run the application program, type **6** "return" when the "Train-Test Program" menu is displayed. Instructions for loading and running the application program will then be displayed. This is discussed in more detail in section 2.3.

2.2.1 Training One Word At A Time

This procedure is used where the teacher is not confident that the student has the strength, concentration or ability to train the full vocabulary set in one sitting. This gives the teacher more control over the training but is more time consuming.

Training in this case is achieved by using the "Train-Test Program" menu as described above in 2.2.1. However, Steps 1 and 2 are not used. After the "Train-Test Program" menu is displayed by using 2.1, then train each word singly by using Step 4 of the description in 2.2.1. This method will probably take several training sessions due to the time taken in selecting words. At the end of each session, use Step 5 to save the newly trained voice patterns. Use Step 3 to determine which sounds lack separability and retrain using Step 4 as required.

2.3 Running the Application Program

The application program is the program which actually performs the required function for the user. Word processing programs, arithmetic practice programs and games programs are application programs.

Use of the Teacher's VIM disk as described in section 2.1 or in section 2.2.1 Step 6, will load to the part of the program which prompts the user in the use of the application program. The prompts displayed by the computer are reprinted in the appendices, Section 4.2 and are discussed in general terms below.

2.3.1 PANIA AND STAR BLAZER

The Teacher's VIM program automatically loads and runs these two programs, no physical contact is required from the user. The prompt supplied by Teacher's VIM remind the user of the command words available and their effect.

2.3.2 BANK STREET WRITER and the EXTENDED NUMBER SET

The Teacher's VIM program prompts the user to remove the Teacher's VIM disk from the disk drive and insert the application program disk and press "control-open apple, reset". This will load the application program but will not affect the voice patterns which are stored in the VIM hardware. Use the application programs in accord with their instruction booklets, but use your voice in place of contact with the keyboard.

2.4 Use in the Classroom

The following arrangements were used in a classroom setting during a study undertaken at the Regency Park Centre for Young Disabled. The study was aimed at determining the suitability of the VIM equipment for several children aged between 7-16 years of age.

Each student was provided with their own box for storage of their Teacher's VIM disks. The Teacher's VIM disks were selected to be appropriate for the student's academic level and were labelled with the student's name.

During the initial introduction period, the teacher demonstrated the use of voice input and the application program with his own set of trained disks. The teacher then generally used the "one word at a time" method to train the student's voice patterns onto the student's disks. In the following lessons the teacher monitored the student's progress and assisted the student in retraining words when required. Copies of the numbered lists in Appendix 4.1 were kept for each student and marked to show how many times words were trained or retrained. This is particularly useful when training by the "one word at a time" method.

Normally 10-20 half hour lessons were required to bring students to the point where a true assessment could be made on the suitability of voice input. It is likely that able bodied users would take less time to become familiar with the equipment, but would be unlikely to use the equipment in practical long term applications.

The following points should be considered when deciding whether a candidate might make practical use of the VIM voice input equipment.

i) The candidate must have good control over a consistent voice. This does not necessarily mean that the candidate's speech would be clear, or even intelligible to the casual listener. On the other hand, clear speakers may be unsuccessful if they do not have the maturity to control the emotional tone in their voice.

ii) The upper limit for rate of entry by voice on the VIM equipment (where one character is entered per utterance) is approximately 30 characters per minute. A candidate able to satisfactorily enter characters by physical means at this, or higher rates is unlikely to wish to use voice input in the long term.

iii) The candidate must be motivated toward, and have a use for a fixed computer system.

3. ADVANCED TOPICS

The subjects discussed in previous sections are sufficient for the student's early contact with the VIM. This will enable the teacher to decide whether the equipment will be of real and lasting value to the student.

If the teacher decides that the student will benefit from the equipment, then it is likely that the teacher will wish to customise the VIM to suit. This section provides information to assist the teacher make these modifications. A working knowledge of the VIM manual and of Applesoft BASIC is assumed.

3.1 Steps 7 to 10 of the "Train-Test Program" Menu

Steps 7,8 and 9 provide access to the VIM subroutines which manipulate the reject threshold and test recognition. Refer to the VIM manual for a description of these procedures.

Step 10 provides access to the main VIM utility menu and all its options. Once the VIM utility menu is entered, then it is not possible to go back to the Teacher's VIM "Train-Test Program" menu. All the functions provided by Teacher's VIM are a subset of the functions provided by the VIM utility program. Refer to the VIM manual for instructions in use of the VIM utility program.

3.2 Vocab and Voice Pattern Files

The VIM utility program allows the use of many different file names and locations for the vocabulary and voice pattern files. However, Teacher's VIM allows only for the use of one file for the vocabulary and one file for voice patterns. They are called VOCAB and VOICE PATTERN respectively. If you wish to modify the contents of these files then use Step 10 on the "Train-Test Program" menu to go to the VIM utility program and use its editing facilities. You will need to refer to the files by the names printed above. If either VOCAB or VOICE PATTERN are erased or damaged, then Teacher's VIM (as supplied) will fail to work properly.

3.3 Hands Off Operation of Software

The equipment and software supplied does not allow for hands off operation of the computer during initial training or retraining. However, hands off operation may be achieved during normal operation.

The supplied Teacher's VIM disks "Pania" and "Star Blazer" load voice patterns and then default through to load and run the games "Pania" and "Star Blazer". These disks allow, for example, the user of a wheelchair to operate the games without physical contact with the keyboard. The disk would be stored in the computer's disk drive. A microswitch would be arranged to detect the presence of a wheelchair in front of the computer monitor, and would turn the power to the computer on. The computer would then load and run the Teacher's VIM disk which would load voice patterns and run the game. After finishing the game the user would move away from the monitor which would cause the microswitch to shut down the computer.

The Bank Street Writer, Extended Number Set disks supplied do not default through to run the application programs. The reason for this may be seen by inspection of the program extracts shown in 4.2.2 and 4.2.3.

The application program Wuzzle (for example) may be arranged to start up automatically at the end of the Teacher's VIM, Extended Number Set disk by :-

i) Placing the Wuzzle program and the sub-programs called by Wuzzle onto the Teacher's VIM disk

ii) Delete lines 60000 to 60090 of the Teacher's VIM HELLO program.

iii) Insert the following lines :-

```
60000    D$ = CHR$(4)
60010    PRINT D$ ; "RUN WUZZLE"
```

APPENDIX A.1.1

Numbered Vocabulary List for Teacher's VIM

Bank Street Writer

Student's name ....

Box number ....

Teacher's name ....

Date ....

01 CORRECTION
02 ESCAPE
03 RETURN
04 EXCLAMATION
05 CURLY A @
06 NUMBER SIGN #
07 DOLLAR SIGN $
08 PERCENTAGE %
09 CAROT ^
10 AMPERSAND &
11 ASTERISK *
12 UNDERLINE _
13 PLUS SIGN +
14 MINUS SIGN -
15 EQUAL SIGN =
16 DELETE
17 ONE
18 TWO
19 THREE
20 FOUR
21 FIVE
22 SIX
23 SEVEN
24 EIGHT
25 NINE
26 ZERO
27 LEFT BRACKET (
28 RIGHT BRACKET)
29 LEFT BRACE {
30 RIGHT BRACE }
31 LEFT SQUARE [
32 RIGHT SQUARE]
33 LESS THAN <
34 GREATER THAN >
35 COLON
36 SEMI-COLON
37 APOSTROPHE '
38 QUOTATION "
39 CRINE `
40 WAVE ~
41 CENTRE LINE |
42 LEFT SLASH \
43 RIGHT SLASH /
44 COMMA ,
45 FULL STOP .
46 QUESTION MARK ?
47 LEFT ARROW
48 RIGHT ARROW
49 UP ARROW
50 DOWN ARROW
51 TABULATE
52 SPACE
53 ALPHA
54 BRAVO
55 CHARLIE
56 DELTAH
57 ECHO
58 FOXTROT
59 GOLF
60 HOTEL
61 INDIA
62 JULIET
63 KILO
64 LIMA
65 MIKE
66 NOVEMBER
67 OSCAR
68 PAPA
69 QUEBEC
70 ROMEO
71 SIERRA
72 TANGO
73 UNIFORM
74 VICTOR
75 WHISKEY
76 XRAY
77 YANKEE
78 ZULU
79 CENTRE
80 INDENT

APPENDIX A.1.2

Numbered Vocabulary List for Teacher's VIM

The Extended Number Set

Student's name ....
Box number ....
Teacher's name ....
Date ....

01 CORRECTION
02 RETURN
03 ZERO
04 ONE
05 TWO
06 THREE
07 FOUR
08 FIVE
09 SIX
10 SEVEN
11 EIGHT
12 NINE
13 YES
14 NO
15 RIGHT
16 LEFT
17 RETURN
18 ZERO
19 ONE
20 TWO
21 THREE
22 FOUR
23 FIVE
24 SIX
25 SEVEN
26 EIGHT
27 NINE
28 YES
29 NO
30 RIGHT
31 LEFT
32 RETURN
33 ZERO
34 ONE
35 TWO
36 THREE
37 FOUR
38 FIVE
39 SIX
40 SEVEN
41 EIGHT
42 NINE
43 YES
44 NO
45 RIGHT
46 LEFT
47 RETURN
48 ZERO
49 ONE
50 TWO
51 THREE
52 FOUR
53 FIVE
54 SIX
55 SEVEN
56 EIGHT
57 NINE
58 YES
59 NO
60 RIGHT
61 LEFT
62 RETURN
63 ZERO
64 ONE
65 TWO
66 THREE
67 FOUR
68 FIVE
69 SIX
70 SEVEN
71 EIGHT
72 NINE
73 YES
74 NO
75 RIGHT
76 LEFT
77 RETURN
78 RETURN
79 RETURN
80 RETURN

APPENDIX A.1.3

Numbered Vocabulary List for Teacher's VIM

Pania

Student's name ....

Box number ....

Teacher's name ....

Date ....

01 CORRECTION
02 DIG
03 UP
04 RIGHT
05 LEFT
06 DOWN
07 STOP
08 FILL
09 START
10 DIG
11 UP
12 RIGHT
13 LEFT
14 DOWN
15 STOP
16 FILL
17 START
18 DIG
19 UP
20 RIGHT
21 LEFT
22 DOWN
23 STOP
24 FILL
25 START
26 DIG
27 UP
28 RIGHT
29 LEFT
30 DOWN
31 STOP
32 FILL
33 START
34 DIG
35 UP
36 RIGHT
37 LEFT
38 DOWN
39 STOP
40 FILL
41 START
42 DIG
43 UP
44 RIGHT
45 LEFT
46 DOWN
47 STOP
48 FILL
49 START
50 DIG
51 UP
52 RIGHT
53 LEFT
54 DOWN
55 STOP
56 FILL
57 START
58 DIG
59 UP
60 RIGHT
61 LEFT
62 DOWN
63 STOP
64 FILL
65 START
66 DIG
67 UP
68 RIGHT
69 LEFT
70 DOWN
71 STOP
72 FILL
73 START
74 DIG
75 UP
76 RIGHT
77 LEFT
78 DOWN
79 STOP
80 FILL

APPENDIX A.1.4

Numbered Vocabulary List for Teacher's VIM

Star Blazer

Student's name ....
Box number ....
Teacher's name ....
Date ....

01 CORRECTION
02 UP
03 DOWN
04 FIRE
05 FASTER
06 SLOWER
07 MAX
08 CRAWL
09 TOP
10 BOTTOM
11 UP
12 DOWN
13 FIRE
14 FASTER
15 SLOWER
16 MAX
17 CRAWL
18 TOP
19 BOTTOM
20 UP
21 DOWN
22 FIRE
23 FASTER
24 SLOWER
25 MAX
26 CRAWL
27 TOP
28 BOTTOM
29 UP
30 DOWN
31 FIRE
32 FASTER
33 SLOWER
34 MAX
35 CRAWL
36 TOP
37 BOTTOM
38 UP
39 DOWN
40 FIRE
41 FASTER
42 SLOWER
43 MAX
44 CRAWL
45 TOP
46 BOTTOM
47 UP
48 DOWN
49 FIRE
50 FASTER
51 SLOWER
52 MAX
53 CRAWL
54 TOP
55 BOTTOM
56 UP
57 DOWN
58 FIRE
59 FASTER
60 SLOWER
61 MAX
62 CRAWL
63 TOP
64 BOTTOM
65 UP
66 DOWN
67 FIRE
68 FASTER
69 SLOWER
70 MAX
71 CRAWL
72 TOP
73 BOTTOM
74 UP
75 DOWN
76 FIRE
77 FASTER
78 SLOWER
79 MAX
80 CRAWL

Application Program Instructions Supplied by Teacher's VIM

A.2.1 Bank Street Writer

THE VOICE INPUT MODULE IS NOW LOADED
WITH THE VOICE PATTERNS FOR BANK STREET
WRITER. PLEASE PLACE THE BANK STREET
WRITER DISC IN THE DISC DRIVE AND PRESS
`CONTROL - OPEN APPLE - RESET'

Screen Prompt

```
60000 PRINT "THE VOICE INPUT MODULE IS NOW LOADED"
60010 PRINT "WITH THE VOICE PATTERNS FOR BANK STREET" 60020
PRINT "WRITER, PLEASE PLACE THE BANK STREET"
60030 PRINT "WRITER DISC IN THE DISC DRIVE AND PRESS"
60040 PRINT "`CONTROL - OPEN APPLE - RESET'"
```

Program lines which generate screen prompt

Application Program Instructions Supplied by Teacher's VIM

A.2.2 Extended Number Set

THE EXTENDED NUMBER SET VOICE PATTERNS HAVE BEEN LOADED, PLACE THE DISC WITH THE APPLICATION PROGRAMME IN THE DISC DRIVE AND PRESS `CONTROL - OPEN APPLE - RESET'.

THIS VOCABULARY IS SUITABLE FOR ;

MATHVADER (FAST AND SLOW)
SMILE
WUZZLE
SPACESHIP

Screen Prompt

```
60000 PRINT "THE EXTENDED NUMBER SET VOICE PATTERNS"
60010 PRINT "HAVE BEEN LOADED, PLACE THE DISC WITH"
60020 PRINT "THE APPLICATION PROGRAMME IN THE DISC"
60030 PRINT "DRIVE AND PRESS `CONTROL - OPEN APPLE -"
60040 PRINT "RESET'."
60045 PRINT
60050 PRINT "THIS VOCABULARY IS SUITABLE FOR ;"
60055 PRINT
60060 PRINT "     MATHVADER (FAST AND SLOW)"
60070 PRINT "     SMILE"
60080 PRINT "     WUZZLE"
60090 PRINT "     SPACESHIP"
```

Program Lines which generate screen prompt

Application Program Instructions Supplied by Teacher's VIM

A.2.3 Pania

```
THE VOICE INPUT MODULE IS NOW LOADED WITH THE
VOICE PATTERNS FOR PANIA.

WAIT WHILE PANIA LOADS THEN SAY `START' TO
START THE GAME.

VOCAB;
      START      STARTS THE GAME
      UP         TO GO UP THE LADDER
      DOWN       TO GO DOWN THE LADDER
      RIGHT      TO GO RIGHT
      LEFT       TO GO LEFT
      DIG        TO DIG A HOLE
      FILL       TO FILL A HOLE
      STOP       TO STOP THE MAN
```

Screen Prompt

```
60000 PRINT "THE VOICE INPUT MODULE IS NOW LOADED"
60010 PRINT "WITH THE VOICE PATTERNS FOR PANIA."
60015 PRINT
60020 PRINT "WAIT WHILE PANIA LOADS THEN SAY `START'"
60030 PRINT "TO START THE GAME."
60040 PRINT
60050 PRINT "VOCAB;"
60060 PRINT "  START      STARTS THE GAME"
60070 PRINT "  UP         TO GO UP THE LADDER"
60080 PRINT "  DOWN       TO GO DOWN THE LADDER"
60090 PRINT "  RIGHT      TO GO RIGHT"
60100 PRINT "  LEFT       TO GO LEFT"
60110 PRINT "  DIG        TO DIG A HOLE"
60120 PRINT "  FILL       TO FILL A HOLE"
60130 PRINT "  STOP       TO STOP THE MAN"
60200 D$ = CHR$ (4)
60210 PRINT D$;"BRUN PANIA"
```

Program lines which generate screen prompt

Application Program Instructions Supplied by Teacher's VIM

A.2.4 Star Blazer

THE VOICE INPUT MODULE IS NOW LOADED WITH THE VOICE PATTERNS FOR STAR BLAZER

WAIT A FEW MOMENTS WHILE STAR BLAZER LOADS, THEN SAY `FIRE' TO START THE GAME

VOCAB;

FIRE	TO START THE GAME
UP	TO GO UP
DOWN	TO GO DOWN
TOP	TO GO UPAT MAXIMUM RATE
BOTTOM	TO GO DOWN AT MAXIMUM RATE
FASTER	TO SPEED UP
SLOWER	TO SLOW DOWN
MAX	MAXIMUM SPEED
CRAWL	MINIMUM SPEED
FIRE	FIRES CANNONS AND DROPS BOMBS

Screen Prompt

```
60000 PRINT "THE VOICE INPUT MODULE IS NOW LOADED"
60010 PRINT "WITH THE VOICE PATTERNS FOR STAR BLAZER"
60018 PRINT
60019 PRINT
60020 PRINT "WAIT A FEW MOMENTS WHILE STAR BLAZER"
60030 PRINT "LOADS, THEN SAY `FIRE' TO START THE GAME"
60042 PRINT
60043 PRINT
60044 PRINT
60050 PRINT "VOCAB;"
60055 PRINT "   FIRE    TO START THE GAME"
60060 PRINT "   UP      TO GO UP"
60070 PRINT "   DOWN    TO GO DOWN"
60080 PRINT "   TOP     TO GO UP AT MAXIMUM RATE"
60090 PRINT "   BOTTOM  TO GO DOWN AT MAXIMUM RATE"
60100 PRINT "   FASTER  TO SPEED UP"
60110 PRINT "   SLOWER  TO SLOW DOWN"
60120 PRINT "   MAX     MAXIMUM SPEED"
60130 PRINT "   CRAWL   MINIMUM SPEED"
60140 PRINT "   FIRE    FIRES CANNONS AND DROPS BOMBS"
60200 D$ = CHR$ (4)
60210 PRINT D$;"BRUN STAR BLAZER"
```

Program lines which generate screen prompt

APPENDIX A3

TEACHER'S VIM

COMPUTER

PROGRAM.

```
]LOAD HELLO
]LIST

1  GOTO 50000
2  HIMEM: (6 * 16 + 15) * 256 + 254: REM $6FFE
5  ONERR  GOTO 20000
10  GOSUB 30000
20  PRINT D$;"BLOAD AVIMCOM,A";RT
30  TEXT : HOME : GOSUB 40000: GOTO 1000
100  REM VIM STATUS
110  POKE OB,24: POKE NB,1: CALL GIVECMD: POKE N2,3: CALL RECEIVE:ST =  FN
    LH(RB): RETURN
200  REM DOWNLOAD VP & VC FILE
210  POKE OB,16: POKE OB + 1, PEEK (RP): POKE OB + 2, PEEK (RP + 1): POKE
    NB,3: CALL DNVPVC
220  POKE OB,4: POKE OB + 1, PEEK (RJ): POKE NB,2: CALL GIVECMD: RETURN
300  REM GET VOCAB FROM VIM
310  GOSUB 100: IF ST = 0 OR ST = 1 THEN  POP : GOTO 20300: REM  NO VOCAB

320  POKE OB,25: POKE OB + 1,0: POKE OB + 2,0: POKE CB + 2,3: POKE N2,6
330  CALL UPLODE: POKE VF + 7,85: POKE VF + 8,170: RETURN
400  REM SET REJECT
405  PRINT : HTAB 4: INPUT "TYPE REJECT THRESHOLD(0 TO 128)";A$: ON  NOT
     LEN (A$) GOTO 420:I =  VAL (A$): IF I < 0 OR I > 128 THEN  PRINT BE$
    : GOTO 400
410  POKE OB,4: POKE OB + 1,I: POKE NB,2: CALL GIVECMD
420  RETURN
500  REM READ REJECT
505  POKE OB,5: POKE NB,1: CALL GIVECMD: POKE N2,1: CALL RECEIVE
510  PRINT : HTAB 4: PRINT "CURRENT REJECT THRESHOLD="; PEEK (RB): GOTO 9
    260
600  REM LOAD
605  GOSUB 9200: IF  NOT  PEEK (IB) THEN E = 1: GOTO 20175
610  GOSUB 9170:DR =  VAL (C$): IF  NOT DR THEN E = 1: GOTO 20175
620  PRINT D$;"BLOAD ";A$;",D";DR;",A";VF:VF$ = A$
630  IF  PEEK (VF + 7) <  > T1 OR  PEEK (VF + 8) <  > T2 THEN  GOSUB 9190
    : GOTO 20140
640  IF T1 = 85 AND T2 = 170 AND  NOT  PEEK (VF + 10) AND  NOT  PEEK (VF +
    11) THEN  CALL COVERT: REM  CONVERT REV 2.0 FILES
650 E = 0
660  ON ZZ GOTO 3920,6930,5110,3140,3500,6000,3820,1505
666  GOTO 20180
675  HTAB 4: PRINT """";A$;"""";" HAS BEEN LOADED": HTAB 4: PRINT "TO THE A
    VIM.": RETURN
700  REM SAVE
702  IF I6 = 1 THEN  GOTO 2200
705  GOSUB 9200: IF  NOT  PEEK (IB) THEN E = 11: RETURN
710  GOSUB 9170:DR =  VAL (C$): IF  NOT DR THEN E = 11: RETURN
720  PRINT D$;"BSAVE ";A$;",A";A;",L";L",D";DR
730  RETURN
800  REM TEST RECOGNITON
802  GOSUB 100: IF ST <  > 257 GOTO 20600
805  HOME : VTAB 10: GOSUB 8600: IF DV% < 0 THEN  RETURN
810  IF DV% AND  NOT PN THEN  RETURN
815  GOSUB 300
820  POKE VF + 6,DV%: IF DV% = 1 THEN  PRINT D$;"PR#";PN
825  GOSUB 9700:WN = 1
830  POKE OB,8: POKE OB + 1,1: POKE OB + 2,252: POKE OB + 3, PEEK (MW): POKE
    OB + 4,255: POKE NB,5: CALL GIVECMD: REM  RECOGNIZE W/#'S
```

```
835  POKE N2,4: CALL TESTVC:I =  PEEK (RB): IF I = 251 THEN  RETURN
840 WN% =  PEEK (RB): IF WN% > 251 THEN  PRINT "***  REJECTED  ***": GOTO
    880
845  POKE CW,WN%: GOSUB 9620: CALL PWRD
850 WN% =  PEEK (RB + 3): POKE CW,WN%: HTAB 19: CALL PWRD
855 I =  PEEK (RB + 1): HTAB 36: PRINT I
860 GS =  PEEK (RB + 2): IF GS < MS THEN MS = GS
865 TS = TS + GS:WC = WC + 1
870  IF I < MD THEN MD = I
875 TD = TD + I
880 WN = WN + 1
885 MM% =  PEEK (MW): IF WN > MM% THEN WN = 1
890  GOTO 835
900  REM SEPERABILITY TEST
902  GOSUB 100: IF ST <  > 257 GOTO 20600
905  HOME : VTAB 10: GOSUB 8600: IF DV% < 0 THEN  RETURN
910  IF DV% AND  NOT PN THEN  RETURN
920  POKE VF + 6,DV%: GOSUB 300
930  IF DV% = 1 THEN  PRINT D$;"PR#";PN
940  GOSUB 9720: POKE OB,26: POKE OB + 1,5: POKE NB,2: CALL GIVECMD
950 DI = 3: POKE N2,4
960  CALL RECEIVE:I =  PEEK (RB): IF I = 255 THEN  PRINT : PRINT : HTAB 1
    0: PRINT "END OF TEST": PRINT D$;"PR#0": GOTO 9260
970 WN% =  PEEK (RB): POKE CW,WN%: GOSUB 9620: CALL PWRD:WN% =  PEEK (RB +
    3): POKE CW,WN%: HTAB 19: CALL PWRD: HTAB 36: PRINT  PEEK (RB + 1);: IF
     PEEK (RB + 1) < DI THEN  HTAB 39: PRINT "X";
980  PRINT : GOTO 960
1000  TEXT
1002  GOTO 1500
1005  HOME : VTAB 3
1010  HTAB 5: PRINT "****************************"
1015  HTAB 5: PRINT "APPLE VOICE UTILITY PROGRAM"
1020  HTAB 5: PRINT "****************************"
1025  PRINT
1030  HTAB 5: PRINT "#            MENU"
1035  HTAB 5: PRINT "-    -------------------------"
1040  HTAB 5: PRINT "1 - BUILD & EDIT VOCABULARY": PRINT ""
1045  HTAB 5: PRINT "2 - TRAIN VOCABULARY": PRINT
1050  HTAB 5: PRINT "3 - RECOGNIZE VOCABULARY": PRINT
1055  HTAB 5: PRINT "4 - TEST VOCABULARY": PRINT
1060  HTAB 5: PRINT "5 - IDLE AVIM": PRINT
1065  HTAB 5: PRINT "6 - QUIT THE VOICE UTILITY": PRINT
1070  HTAB 5: PRINT "-- --------------------------": PRINT
1075  GOSUB 9090
1080  ON MN GOTO 6000,3500,3000,5000,1100,49900
1085  IF MN = 100 THEN  GOSUB 45000
1090  GOTO 1005
1100  REM IDLE
1110  HOME : VTAB 10: HTAB 4: PRINT "THE AVIM WILL BE DEACTIVATED": HTAB
    21: PRINT "-----------"
1120  GOSUB 9270: IF A$ <  > "Y" GOTO 1005
1130  POKE OB,12: POKE OB + 1,1: POKE NB,2: CALL GIVECMD: GOTO 1005
1500  GOTO 2500
1505  TEXT : HOME
1510  HTAB 5: PRINT "******************"
1515  HTAB 5: PRINT "TRAIN-TEST PROGRAM"
1520  HTAB 5: PRINT "******************"
1525  PRINT
1530  PRINT "INITIAL TRAINING PERFORM STEPS 1 TO 6"
```

```
1535  HTAB 5: PRINT "-    ------------------------"
1540  HTAB 5: PRINT "1 - NEW TRAINING"
1545  HTAB 5: PRINT "2 - ADDITIONAL TRAINING (TWICE)"
1550  HTAB 5: PRINT "3 - TEST SEPERABILITY"
1555  HTAB 5: PRINT "4 - RETRAIN SINGLE WORDS LISTED"
1556  HTAB 5: PRINT "IN 3. REPEAT 3 AND 4 UNTIL ALL"
1557  HTAB 5: PRINT "WORDS ARE SEPERABLE. THEN GO TO 5"
1560  HTAB 5: PRINT "5 - SAVE VOICE PATTERNS ON DISC"
1565  HTAB 5: PRINT "6 - RUN APPLICATION PROGRAMME"
1575  HTAB 5: PRINT "7 - READ REJECT THRESHOLD"
1580  HTAB 5: PRINT "8 - SET REJECT THRESHOLD"
1585  HTAB 5: PRINT "9 - TEST RECOGNITION"
1590  HTAB 5: PRINT "10- AVIM UTILITY MENUS"
1592  PRINT
1600  GOSUB 9090
1610  ON MN GOSUB 3600,3700,900,4000,2000,1,500,400,5200
1630  IF MN = 10 THEN  GOTO 1005
1640  GOTO 1505
2000  REM  SAVE TO DISC UNDER FILE NAME "VOICE PATTERN"
2005 I6 = 1
2010  GOSUB 4900
2200  REM  FROM 702
2210 A$ = "VOICE PATTERN"
2220 C$ = "1"
2230 I6 = 0
2240  GOSUB 720
2250  RETURN
2500  REM  LOAD "VOCAB"
2510 ZZ = 8
2520 T1 = 85:T2 = 170:V$ = "V"
2530 A$ = "VOCAB"
2540 DR = 1
2550 VF$ = A$
2600  GOTO 620
3000  HOME : VTAB 3: HTAB 6: PRINT "RECOGNIZE VOCABULARY"
3010  HTAB 6: PRINT "********************"
3015  PRINT : PRINT
3020  HTAB 4: PRINT "#         MENU"
3025  HTAB 4: PRINT "-    ----------------------"
3030  HTAB 4: PRINT "1 - RECOGNIZE VOCABULARY": PRINT
3035  HTAB 4: PRINT "2 - LOAD VOICE PATTERNS &"
3040  HTAB 4: PRINT "    RECOGNIZE VOCABULARY": PRINT
3045  HTAB 4: PRINT "3 - LIST ACTIVE AVIM VOCAB.": PRINT
3050  HTAB 4: PRINT "4 - ";: GOSUB 9210: HTAB 4: PRINT "5 - EXIT"
3055  HTAB 4: PRINT "-    ----------------------": PRINT : GOSUB 9090
3060  ON MN GOTO 3200,3100,3065,3065,1005
3065  ON MN GOSUB ,,7900,8500: GOTO 3000
3100 T1 = 170:T2 = 85:V$ = "R":ZZ = 4: GOTO 600
3140  IF E GOTO 3000
3150  GOSUB 200: HOME : VTAB 10: GOSUB 675: GOSUB 3220
3200  GOSUB 100: IF ST <  > 257 THEN  GOSUB 20600: GOTO 3000
3210  HOME : VTAB 11
3220  POKE OB,21: POKE NB,1: CALL GIVECMD
3230  POKE N2,0: CALL RECEIVE
3240  IF  PEEK (RB) <  > 67 THEN  GOSUB 20200: GOTO 3000
3320  HTAB 5: PRINT "RECOGNITION HAS BEEN ACTIVATED"
3330  PRINT : PRINT " YOU MAY START YOUR APPLICATION PROGRAM"
3340  GOTO 49999
3500  HOME
```

```
3505  HTAB 5: PRINT "*****  TRAIN VOCABULARY  *****": PRINT
3510  HTAB 4: PRINT "#         MENU"
3515  HTAB 4: PRINT "-   ----------------------"
3520  HTAB 4: PRINT "1 - LOAD VOCAB. TO THE AVIM": PRINT
3525  HTAB 4: PRINT "2 - NEW TRAINING": PRINT
3530  HTAB 4: PRINT "3 - ADDITIONAL TRAINING": PRINT
3535  HTAB 4: PRINT "4 - SAVE VOICE PATTERNS ON DISK": PRINT
3540  HTAB 4: PRINT "5 - LOAD VOICE PATTERNS FROM DISK": PRINT
3545  HTAB 4: PRINT "6 - RETRAIN SINGLE WORDS": PRINT
3550  HTAB 4: PRINT "7 - LIST ACTIVE AVIM VOCAB.": PRINT
3555  HTAB 4: PRINT "8 - ";: GOSUB 9210
3560  HTAB 4: PRINT "9 - EXIT"
3565  HTAB 4: PRINT "-   ----------------------": PRINT
3570  GOSUB 9090
3575  IF MN = 1 GOTO 3910
3580  ON MN GOSUB 3910,3600,3700,4900,3800,4000,7900,8500
3585  IF MN = 9 THEN 1090
3590  GOTO 3500
3600 TC = 18: GOSUB 300: IF ST = 257 THEN  GOSUB 9220: IF B$ =  CHR$ (13)
     THEN  RETURN
3610  GOTO 4700
3700 TC = 19: GOSUB 300: ON ST = 257 GOTO 4700: GOTO 20600
3800  REM LOAD VOICE PATTS
3805  POP
3810 T1 = 170:T2 = 85:V$ = "R":ZZ = 7: GOTO 600
3820  IF E GOTO 3500
3830  GOSUB 200: GOSUB 675: GOSUB 9260: GOTO 3500
3900  REM READ FILE FOR TRAIN
3910 T1 = 85:T2 = 170:V$ = "V":ZZ = 1: GOTO 600
3920  IF E GOTO 3500
3930  POKE OB,17:I =  PEEK (MW):I = I + 1: POKE OB + 1,0: POKE OB + 2,I: POKE
    OB + 3,0: POKE OB + 4,0: POKE NB,5: POKE N2,0
3940  CALL DOWNLODE: CALL RECEIVE: IF  PEEK (RB) <  > 67 THEN  GOSUB 2070
    0: GOTO 3590
3970  GOSUB 675: GOSUB 9260: GOTO 3500
4000  REM RETRAIN
4005  GOSUB 300: ON ST <  > 257 GOTO 20600: POKE OB,18
4010  HOME : PRINT : HTAB 10: PRINT "RETRAINING MODES"
4015  HTAB 10: PRINT "***************"
4020  PRINT
4025  HTAB 4: PRINT "#        MENU"
4030  HTAB 4: PRINT "-   ------------------------"
4035  HTAB 4: PRINT "TRAIN ONE WORD BY :": PRINT
4040  HTAB 4: PRINT "1 - SPELLING": PRINT
4045  HTAB 4: PRINT "2 - INDEX NUMBER": PRINT
4050  HTAB 4: PRINT "3 - LIST ACTIVE AVIM VOCAB.": PRINT
4055  HTAB 4: PRINT "4 - EXIT"
4060  HTAB 4: PRINT "-   ------------------------": PRINT
4065  GOSUB 9090
4070  ON MN GOTO 4100,4200
4075  IF MN = 3 THEN  GOSUB 8000: GOTO 4010
4080  IF MN = 4 THEN  RETURN
4085  GOTO 4010
4100  GOSUB 9000: PRINT "TRAINING VOCABULARY, OR": GOSUB 9020: PRINT : HTAB
    4: INPUT A$
4110 I =  PEEK (IB): IF I = 0 THEN  RETURN
4120  CALL MI: POKE OB + 1,2: CALL GIVECMD: POKE N2,0: CALL RECEIVE
4130 A$ =  CHR$ ( PEEK (RB)): IF A$ = "E" GOTO 20500
4140  GOSUB 4820:XT = 3: GOSUB 7570: GOTO 4100
```

```
4200  GOSUB 9000: PRINT "WORD NUMBER, (1-"; PEEK (MW);") OR": GOSUB 9020:
     INPUT A$
4210 I =  VAL (A$): IF  NOT (I) THEN  RETURN
4220 J =  PEEK (MW): IF I > J GOTO 20400
4230  POKE OB + 1,3: POKE OB + 2,I: POKE OB + 3,I: POKE OB + 4,255: POKE
    NB,5
4240  POKE N2,0: CALL GIVECMD: CALL RECEIVE:A$ =  CHR$ ( PEEK (RB)): IF A
    $ = "E" GOTO 20500
4250  GOSUB 4820:XT = 3: GOSUB 7570: GOTO 4200
4700  REM  TRAIN ALL TC=18 NEW , =19 UPDATE
4702  HOME : VTAB 5: PRINT : HTAB 5: IF TC = 19 THEN  PRINT "ADDITIONAL "
    ;
4704  IF TC = 18 THEN  PRINT "NEW ";
4705  PRINT "TRAINING": GOSUB 9230: IF B$ =  CHR$ (13) THEN  RETURN
4710  POKE OB,TC: POKE OB + 1,1: POKE OB + 2,255: POKE NB,3: CALL GIVECMD
    : GOSUB 4800
4720  IF TC = 18 THEN I = 106: GOSUB 410
4730  GOTO 9260
4800  REM TRAIN PROMPTS
4810  POKE N2,0: CALL RECEIVE:A$ =  CHR$ ( PEEK (RB))
4820  HOME : VTAB 5: CALL 64680: IF A$ = "C" THEN  HTAB 5: PRINT "SAY ...
    ": VTAB 12: HTAB 15: CALL PV: GOTO 4800
4830  IF A$ = "E" THEN  HTAB 2: PRINT "SAY AGAIN ...": VTAB 12: HTAB 15: CALL
    PV: GOTO 4800
4840  HOME : VTAB 12: HTAB 7: INVERSE : PRINT "!!";: NORMAL : PRINT " END
     OF THE TRAINING ";: INVERSE : PRINT "!!": NORMAL : GOSUB 7570: RETURN

4900  REM SAVE VP&VC
4905  GOSUB 100: ON ST <  > 257 GOTO 20600:ZZ = 5
4910 VF$ = "": POKE OB,20: POKE OB + 1,0: POKE OB + 2,0: POKE NB,3
4920  CALL LODEVPVC:V$ = "R": POKE VF + 7,170: POKE VF + 8,85
4930  POKE OB,5: POKE NB,1: CALL GIVECMD: POKE N2,1: CALL RECEIVE: POKE R
    J, PEEK (RB)
4940 A = RP:L =  FN HL(RP) + 32
4950 E = 0: GOSUB 700: PRINT : IF E THEN  RETURN
4960  PRINT : HTAB 4: PRINT "THE VOICE PATTERNS & VOCABULARY": HTAB 4: PRINT
    "HAVE BEEN SAVED ON DISK."
4970  GOTO 9260
5000  HOME : VTAB 3
5005  HTAB 7: PRINT "TEST VOCABULARY"
5010  HTAB 7: PRINT "***************": PRINT
5015  HTAB 4: PRINT "#       MENU"
5020  HTAB 4: PRINT "-    --------------------"
5025  HTAB 4: PRINT "1 - TEST RECOGNITION": PRINT
5030  HTAB 4: PRINT "2 - LOAD VOICE PATTERNS &": HTAB 4: PRINT "     TEST
    RECOGNITION": PRINT
5035  HTAB 4: PRINT "3 - TEST SEPARABILITY": PRINT
5040  HTAB 4: PRINT "4 - READ REJECT THRESHOLD": PRINT
5045  HTAB 4: PRINT "5 - SET REJECT THRESHOLD": PRINT
5050  HTAB 4: PRINT "6 - EXIT"
5055  HTAB 4: PRINT "-    ---------------------": PRINT
5060  GOSUB 9090
5065  IF MN = 2 GOTO 5100
5070  ON MN GOSUB 5200,5100,900,500,400
5075  IF MN = 6 THEN 1090
5080  GOTO 5000
5100 T1 = 170:T2 = 85:V$ = "R":ZZ = 3: GOTO 600
5110  IF E GOTO 5000
5120  GOSUB 200: HOME : VTAB 6: GOSUB 675: GOSUB 5200: GOTO 5000
```

```
5200 TS = 0:WC = 0:MS = 128:TD = 0:MD = 128: GOSUB 800
5210  IF WC = 0 THEN  RETURN
5220 AS% = TS / WC
5230  PRINT : PRINT : HTAB 4: PRINT "MINIMUM SCORE=";MS
5240  PRINT : HTAB 4: PRINT "AVERAGE SCORE=";AS%
5250  PRINT : HTAB 4: PRINT "MINIMUM DIFFERENCE=";MD
5260 TD% = TD / WC: PRINT : HTAB 4: PRINT "AVERAGE DIFFERENCE=";TD%
5270  PRINT : HTAB 4: PRINT "TOTAL TRIES=";WC
5280  PRINT : PRINT D$;"PR#0": GOSUB 9260: RETURN
6000  REM  BUILD & EDIT MODE
6005  HOME : PRINT
6010  HTAB 4: PRINT "BUILD & EDIT VOCABULARY"
6015  HTAB 4: PRINT "***********************": PRINT
6020  HTAB 4: PRINT "#         MENU"
6025  HTAB 4: PRINT "-    --------------------"
6030  HTAB 4: PRINT "1 - BUILD NEW VOCABULARY": PRINT
6035  HTAB 4: PRINT "2 - LIST VOCABULARY": PRINT
6040  HTAB 4: PRINT "3 - EDIT VOCABULARY": PRINT
6045  HTAB 4: PRINT "4 - SAVE VOCAB. ON DISK": PRINT
6050  HTAB 4: PRINT "5 - READ VOCAB. FROM DISK": PRINT
6055  HTAB 4: PRINT "6 - ";: GOSUB 9210
6060  HTAB 4: PRINT "7 - EXIT"
6065  HTAB 4: PRINT "-    ----------------------"
6070  PRINT : GOSUB 9090
6075  IF MN = 5 GOTO 6900
6080  ON MN GOSUB 6200,8000,7000,6500,6900,8500
6085  TEXT : IF MN = 7 THEN 1090
6090  GOTO 6000
6200  PRINT "": HTAB 4: INPUT "NEW VOCABULARY REALLY (Y/N) ";A$: IF A$ <
     > "Y" THEN  RETURN
6210 VF$ = "": CALL CLRMEM: CALL NEVOCAB: POKE VF + 7,85: POKE VF + 8,170
    : GOSUB 9635:WN% =  PEEK (MW): POKE CW,WN%: GOSUB 9610:WN% = WN% + 1:
     POKE CW,WN%: GOTO 7230
6500  IF  PEEK (MW) = 0 THEN A$ = "SAVED": GOTO 9250
6510  IF  FN HL(VF + 7) <  > 21930 GOTO 20500
6520 V$ = "VOCABULARY":A = VF:L =  FN LH(VF + 3) - VF + 32:ZZ = 6: GOTO 7
    00
6900  REM READ FILE FOR BUILD & EDIT
6910 T1 = 85:T2 = 170:V$ = "V":ZZ = 2
6920  GOTO 600
6930  IF E GOTO 6000
6940  GOSUB 675: GOSUB 9260: GOTO 6000
7000  REM EDIT VOCABULARY
7002  IF  PEEK (MW) = 0 OR  FN HL(VF + 7) <  > 21930 THEN A$ = "MODIFIED"
    : GOTO 9250
7005  TEXT : HOME : PRINT : HTAB 8: PRINT "EDIT VOCABULARY"
7010  HTAB 8: PRINT "***************": PRINT
7015  HTAB 4: PRINT "#         MENU"
7020  HTAB 4: PRINT "-    --------------------"
7025  HTAB 4: PRINT "1 - ADD NEW VOCABULARY"
7030  PRINT
7035  HTAB 4: PRINT "2 - CORRECT VOCABULARY"
7040  PRINT
7045  HTAB 4: PRINT "3 - INSERT NEW VOCABULARY"
7050  PRINT
7055  HTAB 4: PRINT "4 - DELETE VOCABULARY"
7060  PRINT : HTAB 4: PRINT "5 - LIST VOCABULARY"
7065  PRINT : HTAB 4: PRINT "6 - EXIT"
```

```
7070  HTAB 4: PRINT "-   ----------------------": PRINT
7075  GOSUB 19020: GOSUB 9090: ON MN GOSUB 7200,7300,7400,7500,8000
7080  TEXT : IF MN = 6 THEN  RETURN
7085  GOTO 7000
7200  REM ADD
7205 WN% =  PEEK (MW):WN% = WN% + 1: POKE CW,WN%: GOSUB 9635
7210  IF  FN MF(0) < 0 GOTO 19010
7220  IF  FN MF(0) < 68 GOTO 19000
7230  GOSUB 19020: POKE 35,23: GOSUB 9620: CALL ADDVC:FG =  PEEK (EDXIT):
     IF FG = 1 GOTO 9570
7240  GOSUB 9610:WN% = WN% + 1: POKE CW,WN%: IF WN% = 100 GOTO 9570
7250  GOTO 7210
7300  REM CHANGE
7305  GOSUB 9635
7310  GOSUB 19020: POKE 35,23:A$ = "CORRECT": VTAB 10: GOSUB 9480: HOME :
     IF  NOT  LEN (A$) THEN  RETURN
7320 WN% =  VAL (A$): POKE CW,WN%:MM% =  PEEK (MW): IF WN% > MM% GOTO 204
    00
7330  HOME : GOSUB 9620: CALL CHANGECW: GOSUB 9610: IF  FN MF(0) < 0 GOTO
    19010
7340  GOTO 7310
7400  REM INSERT
7405  HOME :A$ = "INSERT": GOSUB 9640: VTAB 7: GOSUB 9480: HOME : IF  NOT
     LEN (A$) THEN  RETURN
7410 WN% =  VAL (A$): POKE CW,WN%:MM% =  PEEK (MW): IF WN% > MM% GOTO 204
    00
7420  IF  FN MF(0) < 0 GOTO 19010
7430  IF  FN MF(0) < 92 GOTO 19000
7440  GOSUB 19020: POKE 35,23: GOSUB 9620: CALL ISERTCW: IF  PEEK (EDXIT)
     = 1 GOTO 9570
7450  POKE CW,WN%: GOSUB 9610:WN% = WN% + 1: POKE CW,WN%: GOTO 7420
7500  REM DELETE
7505  PRINT : GOSUB 8800: IF EF% = 1 THEN  RETURN
7510  GOSUB 8900: PRINT : PRINT : HTAB 4: PRINT "THE ABOVE LISTED VOCABUL
    ARY WILL BE": HTAB 4: PRINT "DELETED": PRINT
7515  PRINT : INPUT " DO YOU REALLY WANT IT DELETED (Y/N) ";A$: IF A$ <  >
    "Y" THEN  RETURN
7520  FOR XC = EV% TO BV% STEP  - 1: POKE CW,XC: CALL DLTECW: NEXT
7530  PRINT : HTAB 4: PRINT "VOCAB. NUMBER ";BV%;
7540  IF BV% <  > EV% THEN  PRINT " TO ";EV%;" ARE DELETED"
7550  IF BV% = EV% THEN  PRINT " IS DELETED"
7560 XT = 20: GOSUB 7570: RETURN
7570  FOR XC = 1 TO XT: CALL 64680: NEXT XC:XC =  FRE (0): RETURN
7900  REM LIST ACTIVE VOCAB.
7910  GOSUB 300: GOSUB 8000: POKE MW,0: RETURN
8000  REM LIST VOCABULARY
8002  IF  PEEK (MW) = 0 OR  FN HL(VF + 7) <  > 21930 THEN A$ = "LISTED": GOTO
    9250
8005  HOME : PRINT : HTAB 13: PRINT "LIST VOCABULARY"
8010  PRINT : PRINT "FILENAME: ";VF$
8015  PRINT : PRINT
8020  HTAB 4: PRINT "#        MENU"
8025  HTAB 4: PRINT "--   ------------------------"
8030  HTAB 4: PRINT "1 - ALL VOCAB. ON THE MONITOR": PRINT
8035  HTAB 4: PRINT "2 - ALL VOCAB. ON THE PRINTER": PRINT
8040  HTAB 4: PRINT "3 - ANY VOCAB. ON THE MONITOR": PRINT
8045  HTAB 4: PRINT "4 - ANY VOCAB. ON THE PRINTER": PRINT
8050  HTAB 4: PRINT "5 - EXIT": PRINT
8055  HTAB 4: PRINT "--   ------------------------"
```

```
8060  GOSUB 19020: PRINT : GOSUB 9090: IF  NOT MN GOTO 8000
8065  ON MN GOSUB 8100,8300,8110,8400: TEXT : RETURN
8100 WN% = 1:MM% =  PEEK (MW): GOTO 8120
8110  GOSUB 8800: IF EF% = 1 THEN  RETURN
8120  GOSUB 8900: PRINT : HTAB 4: PRINT "‼ END OF VOCABULARY ‼": GOTO 9
    260
8200 WN% = 1: POKE CW,1:MM% =  PEEK (MW): GOTO 8220
8210  GOSUB 8800: POKE CW,WN%: IF EF% = 1 THEN  RETURN
8220  HOME : POKE DEVICE,1: PRINT D$;"PR#";PN: GOSUB 9640
8230  GOSUB 9360:WN% = WN% + 1: POKE CW,WN%
8240  IF WN% > MM% THEN  HTAB 4: PRINT "‼ END OF VOCABULARY ‼": PRINT :
     PRINT : PRINT : PRINT D$;"PR#0": RETURN
8250  GOTO 8230
8300  IF  NOT (PN) THEN  GOSUB 8640: IF  NOT (PN) THEN  RETURN
8310  GOTO 8200
8400  IF  NOT (PN) THEN  GOSUB 8640: IF  NOT PN THEN  RETURN
8410  GOTO 8210
8500  REM LIST FILENAMES ON DISK
8505  HOME : VTAB 10: GOSUB 9210:S =  PEEK (43626):DR =  PEEK (43624): PRINT
    "  IN SLOT ";S;"   DRIVE ";DR;: INPUT " (Y/N) ";A$: IF A$ = "Y" GOTO
    8520
8510  GOSUB 9180: ON  NOT  LEN (CO$) OR  NOT  VAL (CO$) GOTO 8530: GOSUB
    9170: ON  NOT  LEN (C$) OR  NOT  VAL (C$) GOTO 8530:DR =  VAL (C$):S =
     VAL (CO$)
8520  PRINT D$;"CATALOG,D";DR;",S";S: GOTO 9260
8530  RETURN
8600  REM SELECT OUTPUT DEVICE
8605  HTAB 4: PRINT "TYPE PROMPT DEVICE NUMBER": HTAB 4: PRINT "  (MONITO
    R=1 PRINTER=2), OR": GOSUB 9020: INPUT A$
8610 DV% =  VAL (A$) - 1: IF DV% > 1 OR DV% < 0 THEN DV% =  - 1: RETURN
8620  IF  NOT DV% THEN  RETURN
8630  IF PN THEN  RETURN
8640  PRINT : CALL  - 958: HTAB 4: INPUT "DO YOU HAVE A PRINTER (Y/N) ";A
    $: IF A$ <  > "Y" THEN  RETURN
8650  PRINT : HTAB 4: INPUT "WHAT SLOT IS YOUR PRINTER IN (1-7) ";A$:PN =
     VAL (A$): ON PN > 7 OR PN < 0 GOTO 8650: RETURN
8800  REM SELECT WORD RANGE
8805 MM% =  PEEK (MW): GOSUB 9000: PRINT : HTAB 5: PRINT "STARTING VOCAB.
     NUMBER (1-";MM%;") ";: INPUT B$:BV% =  VAL (B$): IF BV% < 1 THEN EF%
     = 1: RETURN
8810  GOSUB 9010: PRINT : HTAB 5: PRINT "ENDING VOCAB. NUMBER (";BV%;"-";
    MM%;") ";: INPUT B$:EV% =  VAL (B$)
8820  IF BV% > MM% OR BV% < 1 OR EV% > MM% OR EV% < 1 THEN EF% = 1: GOTO
    20400
8840  IF BV% > EV% THEN EF% = 1: GOTO 20800
8870 EF% = 0:WN% = BV%:MM% = EV%: RETURN
8900 TN% = WN% + 7: POKE CW,WN%: GOSUB 9635
8910  GOSUB 9610:WN% = WN% + 1: POKE CW,WN%
8920  IF WN% > MM% THEN  RETURN
8930  IF WN% > TN% THEN TN% = TN% + 9: HTAB 4: INVERSE : PRINT "HIT ANY K
    EY TO CONTINUE LISTING";: NORMAL : GET A$
8940  GOTO 8910
9000  HOME : VTAB 7
9010  PRINT : HTAB 4: INVERSE : PRINT "TYPE";: NORMAL : PRINT " ";: RETURN

9020  HTAB 4: PRINT "DEPRESS 'RETURN' KEY TO ABORT ";: RETURN
9090  HTAB 4: INVERSE : PRINT "TYPE";: NORMAL : INPUT " MENU NUMBER ";A$:
    MN =  ABS ( VAL ( LEFT$ (A$ + "    ",3))): ON MN > 254 GOTO 9090: RETURN
```

```
9170  GOSUB 9010: PRINT "DISK DRIVE NUMBER, OR": GOSUB 9020: PRINT : PRINT
    : HTAB 5:DR =  PEEK (43624): PRINT DR;: HTAB 4: INPUT C$: ON  LEN (C$
    ) AND ( VAL (C$) > 2 OR  VAL (C$) < 1) GOTO 9170: RETURN
9180  GOSUB 9010: PRINT "DISK SLOT NUMBER, OR": GOSUB 9020: PRINT : PRINT
    : HTAB 5: PRINT S;: HTAB 4: INPUT CO$: ON  VAL (CO$) > 7 OR  VAL (CO$
    ) < 0 GOTO 9180: RETURN
9190  GOSUB 9580: PRINT : HTAB 4: INVERSE : PRINT "!! ERROR";: NORMAL : PRINT
    " - ";: RETURN
9200  GOSUB 9000: PRINT "FILE NAME ON THE DISK, OR": GOSUB 9020: PRINT : PRINT
    : HTAB 5: PRINT VF$;: HTAB 4: INPUT A$: RETURN
9210  PRINT "LIST FILE NAMES ON DISK": PRINT : RETURN
9220  HOME : VTAB 5: PRINT "  CURRENT VOICE PATTERNS WILL BE LOST": GOSUB
    9580
9230  VTAB 7: PRINT : HTAB 4: PRINT "PRESS ANY KEY TO CONTINUE OR,": PRINT
    : GOSUB 9020: GET B$: PRINT B$: RETURN
9250  GOSUB 9190: PRINT "NO VOCAB. TO BE ";A$
9260  INVERSE : PRINT : HTAB 6: PRINT " HIT ANY KEY TO GET MENU BACK ";: NORMAL
    : GET B$: RETURN
9270  HTAB 4: INPUT "ARE YOU SURE (Y/N) ";A$: RETURN
9360  GOSUB 9620: CALL S2ETCW: CALL PRNTVC: PRINT : RETURN
9480  GOSUB 9010: PRINT "VOCABULARY NUMBER (1-"; PEEK (MW);")": HTAB 10: PRINT
    " TO ";A$;", OR": GOSUB 9020: INPUT A$: RETURN
9570  PRINT : HTAB 4: PRINT "!! END OF EDITING !!": GOTO 9260
9580  CALL  - 1059: CALL  - 1059: RETURN
9610  HTAB 1: GOSUB 9620: CALL S2ETCW: POKE DEVICE,0: CALL PRNTVC: PRINT
    : PRINT : RETURN
9620  IF WN% < 10 THEN  PRINT "0";
9630  PRINT WN%;" ";: RETURN
9635  HOME
9640  PRINT "FILENAME: ";VF$: PRINT : PRINT "-- ---------------   -------
    --------"
9650  PRINT "#  TRAINING VOCAB.  ;KEY REPLACEMENTS"
9660  PRINT "-- ---------------   ---------------": POKE 34, PEEK (37): RETURN

9700  HOME : VTAB 2: HTAB 10: PRINT "RECOGNITION TEST": HTAB 10: PRINT "-
    ----------------"
9710  PRINT : PRINT : PRINT " #    RECOGNIZED     NEAREST       DIFF": GOTO
    9740
9720  HOME : VTAB 2: HTAB 12: PRINT "SEPARABILITY TEST": HTAB 12: PRINT "
    -----------------"
9730  PRINT : PRINT : PRINT " #     REFERENCE     NEAREST       DIFF"
9740  PRINT "          WORD        NEIGHBOR      ERENCE"
9750  PRINT "-- -------------- --------------   ----": RETURN
19000  PRINT : PRINT BE$;"  VOCABULARY  SPACE  FULL": PRINT "  NO ROOM FO
    R ADDITIONAL WORDS": GOSUB 7560: GOTO 9570
19010  PRINT : PRINT BE$;" OVERFLOW  VOCABULARY TOO LONG ! ": PRINT " PLE
    ASE DELETE "; ABS ( FN MF(0));" CHARACTERS ": GOSUB 7560: GOTO 7330
19020 V =  PEEK (37):H =  PEEK (36): VTAB 24: HTAB 1: IF  FN MF(92) < 0 THEN
     INVERSE : PRINT "NO";: NORMAL
19030  PRINT " ROOM FOR ADDITIONAL VOCABULARY";: HTAB H + 1: VTAB V + 1: PRINT
    ;: RETURN
20000  REM ERROR TRAP
20010 EL =  FN LH(218): REM LINE OF ERROR
20020 ER =  PEEK (222): POKE 216,0: IF ER > 15 THEN  PRINT "ERROR # ";ER;
    " LINE ";EL: END
20025 E = ER: GOSUB 9190: PRINT " DISK ";
20030  IF ER = 6 THEN  PRINT "'FILE NOT FOUND'"
20040  IF ER = 4 THEN  PRINT "'WRITE PROTECTED'": PRINT : HTAB 4: PRINT "
```

```
    PLEASE REMOVE TAB OR": HTAB 7: PRINT "INSERT A DIFFERENT DISK"
20045  IF ER = 8 THEN  PRINT "'I/O ERROR'"
20050  IF ER = 10 THEN  PRINT "'FILE LOCKED'"
20060  IF ER = 13 THEN  PRINT "'FILE TYPE MISMATCH'": PRINT : HTAB 4: PRINT
    "FILENAME USED BY ANOTHER FILE": HTAB 4: GOTO 20140
20080  IF ER = 9 THEN  PRINT "FULL"
20090  IF ER = 11 THEN  PRINT "'SYNTAX ERROR'": PRINT : HTAB 4: PRINT "IL
    LEGAL FILENAME"
20099  GOTO 20170
20140 E = 257
20150  IF V$ = "V" THEN  PRINT "IT IS NOT A VOCAB. FILE": GOTO 20170
20160  PRINT "IT IS NOT A VOICE PATTERNS": HTAB 15: PRINT "FILE.": GOTO 2
    0170
20170  POKE 216,0: ONERR  GOTO 20000
20172  GOSUB 9260
20175  ON ZZ GOTO 3920,6930,5110,3140,3500,6000,3820
20180  PRINT "FATAL SOFTWARE ERROR! CONTACT VMC.": END
20200  GOSUB 9190: PRINT "COMMAND CANN'T BE": HTAB 15: PRINT "ACCEPTED BY
     THE AVIM.": GOTO 9260
20300  GOSUB 9190: PRINT "NO VOCAB. IN THE AVIM.": GOTO 9260
20400  GOSUB 9190: PRINT "THE VOCAB. NUMBER IS": HTAB 15: PRINT "OUT OF R
    ANGE.": GOTO 9260
20500  GOSUB 9190: PRINT "VOCAB. CAN NOT BE FOUND": GOTO 9260
20600  GOSUB 9190: PRINT "NO VOICE PATTERNS IN": HTAB 15: PRINT "THE AVIM
    ": GOTO 9260
20700  GOSUB 9190: PRINT "THE VOCAB. FILE CAN NOT BE": HTAB 15: PRINT "LO
    ADED TO THE AVIM": GOTO 9260
20800  GOSUB 9190: PRINT "THE STARTING IS GREATER ": HTAB 15: PRINT "THAN
     THE ENDING NUMBER.": GOTO 9260
30000  REM  MACHINE CALLS
30005 RT = (7 * 16) * 256: REM SUB-ROUTINE ENTRY($7000)
30010 CLRMEM = RT
30015 S2ETCW = RT + 3: REM SET TO CURRENT WORD
30020 UPLODE = RT + 6
30025 INIT = RT + 9
30030 NEVOCAB = RT + 12
30035 VF = (7 * 16 + 9) * 256
30040 PRNTVC = RT + 15
30045 ADDVC = RT + 18
30050 DLTECW = RT + 21
30055 RP = VF: REM REP. PATTERN
30060 ISERTCW = RT + 24
30065 DNVPVC = RT + 27
30070 LODEVPVC = RT + 30
30075 DOWNLODE = RT + 33
30080 IB = 512
30085 GIVECMD = RT + 36
30090 RECEIVE = RT + 39
30095 PV = RT + 42: REM PRINT RECOGNITION
30100 MI = RT + 45: REM MOVE INPUT
30105 CB = (7 * 16 + 8) * 256: REM AVIM BUFF(7800)
30110 OKAY = CB + 1: REM 0=GOOD
30115 TESTVC = RT + 48
30120 N2 = CB + 3: REM #BYTES RESPONSE
30125 MW = VF:CW = MW + 1:EDXIT = VF + 5:DEVICE = VF + 6: REM MW= # OF WO
    RDS CW= CURRENT WORD
30130 CF = VF + 10: REM VOCAB CHAR COUNT LOW,HI
30135 RJ = VF + 9: REM  REJECT THRESHOLD
30140 PWRD = RT + 51: REM PRINT SPOKEN
```

```
30145 COVERT = RT + 57: REM CONVERTOLD - NEW
30150 CHANGECW = RT + 54
30155 SN = CB:NB = CB + 2:OB = CB + 32:RB = CB + 160
30160 MEMSIZE = (80 * 67) + 1600
30165 D$ =  CHR$ (4):BE$ =  CHR$ (7)
30170 S =  PEEK (43626):DR =  PEEK (43624)
30175  DEF  FN LH(I) =  PEEK (I) +  PEEK (I + 1) * 256: DEF  FN MF(I) = M
    EMSIZE -  FN LH(CF) - (72 *  PEEK (MW)) - I - 130: DEF  FN HL(I) =  PEEK
    (I) * 256 +  PEEK (I + 1)
30180 PN = 0
30185  RETURN
40000  PRINT D$;"BLOAD SLOT,A";SN
40010  CALL INIT:I =  PEEK (OKAY)
40020  IF I = 0 THEN  RETURN
40030  GOSUB 9190: VTAB 10: HTAB 4: PRINT "IF YOUR AVIM IS NOT "
40040  HTAB 4: PRINT "PLUGGED-IN AT APPLE I/O SLOT ";( PEEK (SN) - 128) /
    16
40050  PRINT : HTAB 4: PRINT "PLEASE TYPE THE CURRENT SLOT": HTAB 4: PRINT
    "NUMBER";
40060  INPUT A$:I =  VAL (A$): IF I < 1 OR I > 7 GOTO 40050
40070 I = 128 + I * 16: POKE CB,I: CALL INIT: IF  PEEK (OKAY) GOTO 40030
40080  PRINT D$;"BSAVE SLOT,A";SN;",L8"
40090  RETURN
45000  HOME : VTAB 4
45005  HTAB 7: PRINT "SET/READ PARAMETERS"
45010  HTAB 7: PRINT "*******************": PRINT ""
45015  HTAB 4: PRINT " #        MENU"
45020  HTAB 4: PRINT "--   --------------------"
45025  HTAB 4: PRINT "1 - SET REJECT THRESHOLD": PRINT
45030  HTAB 4: PRINT "2 - READ REJECT THRESHOLD": PRINT
45035  HTAB 4: PRINT "3 - SET WORD PARAMETERS": PRINT
45040  HTAB 4: PRINT "4 - READ WORD PARAMETERS": PRINT
45045  HTAB 4: PRINT "5 - EXIT"
45050  HTAB 4: PRINT "--  --------------------"
45055  PRINT : GOSUB 9090
45060  ON MN GOSUB 400,500,45300,45500
45065  IF MN = 5 THEN  RETURN
45070  GOTO 45000
45300  HTAB 4: INPUT "TYPE T1=";I: POKE OB,9: POKE OB + 1,I: POKE NB,7
45310  HTAB 4: INPUT "TYPE T2=";I: POKE OB + 2,I
45320  HTAB 4: INPUT "TYPE ETHL=";I: POKE OB + 3,I
45330  HTAB 4: INPUT "TYPE MIN. SAMPLE NUMBER=";I: POKE OB + 4,I: HTAB 4:
     INPUT "TYPE NOISE SAMPLE NUMBER=";I: POKE OB + 5,I
45340  HTAB 4: INPUT "TYPE NOISE THRESHOLD=";I: POKE OB + 6,I: CALL GIVEC
    MD: RETURN
45500  POKE OB,10: POKE NB,1: CALL GIVECMD: POKE N2,6: CALL RECEIVE: HOME
    : VTAB 6
45510  HTAB 4: PRINT "T1= "; PEEK (RB): PRINT : HTAB 4: PRINT "T2= "; PEEK
    (RB + 1): PRINT
45520  HTAB 4: PRINT "ETHL= "; PEEK (RB + 2): PRINT : HTAB 4: PRINT "MIN.
     SAMPLE= "; PEEK (RB + 3)
45530  PRINT : HTAB 4: PRINT "NOISE SAMPLE NUMBER= "; PEEK (RB + 4): PRINT
    : HTAB 4: PRINT "NOISE THRESHOLD= "; PEEK (RB + 5): GOTO 9260
49900  TEXT : HOME
49999  POKE 216,0: GOTO 60000
50000  REM CHANGE VOCABULARY PROGRAM, VERSIOIN 2.1,4/11/83
50002  ONERR  GOTO 50082
50004  TEXT : HOME
50006  GOSUB 50104
```

```
50008 D$ =  CHR$ (4): PRINT D$;"BLOAD AVIMCOM"
50010  PRINT D$;"BLOAD SLOT,A";SN
50012  CALL INIT
50014  REM  QUICK LOAD AVIM PROGRAM
50016 T1 = 170:T2 = 85:V$ = "R": GOTO 50054
50018  IF E = 1 THEN  GOTO 50016
50020  GOSUB 50022: GOTO 50026
50022  POKE OB,16:I =  PEEK (RP): POKE OB + 1,I:I =  PEEK (RP + 1): POKE
    OB + 2,I: POKE NB,3
50024  CALL UPLODE: RETURN
50026  HOME : VTAB 10: GOSUB 50044
50028 I = 106
50030  POKE OB,4: POKE OB + 1,I: POKE NB,2: CALL GIVE: GOTO 50036
50032  GOSUB 50050:I =  PEEK (RB): IF I = 0 THEN  GOSUB 50078: GOTO 50014

50034  HOME : VTAB 10: GOSUB 50046
50036  POKE OB,21: POKE NB,1: CALL GIVE
50038  POKE CB + 3,0: CALL RECEIVE
50040 I =  PEEK (RB): IF I <  > 67 THEN  GOSUB 50076: GOTO 50014
50042  HOME : GOTO 53000
50044  HTAB 4: PRINT "";A$;"";" HAS BEEN LOADED": HTAB 4: PRINT "TO THE
     AVIM.": GOTO 50048
50046  HTAB 4: PRINT "RECOGNITION HAS BEEN ACTIVATED."
50048 A$ = "A": RETURN
50050  POKE OB,24: POKE NB,1: CALL GIVE
50052  POKE CB + 3,3: CALL RECEIVE: RETURN
50054  GOSUB 50094:I =  PEEK (512): IF I = 0 THEN E = 1: RETURN
50056  GOSUB 50090:DR =  VAL (C$): IF DR = 0 THEN E = 1: RETURN
50058 D$ =  CHR$ (4): PRINT D$;"BLOAD ";A$;",D";DR;",A";VF
50060 TY =  PEEK (VF + 7): IF TY <  > T1 THEN  GOTO 50066
50062 TY =  PEEK (VF + 8): IF TY <  > T2 THEN  GOTO 50066
50064 E = 0: GOTO 50018
50066 E = 1
50068  GOTO 50080
50070  HOME : VTAB 7
50072  PRINT "": HTAB 4: INVERSE : PRINT "TYPE";: NORMAL : PRINT " ";: RETURN

50074  HTAB 4: PRINT "DEPRESS 'RETURN' KEY TO ABORT ";: RETURN
50076  GOSUB 50092: PRINT "COMMAND CANN'T BE": GOSUB 50100: PRINT "ACCEPT
    ED BY THE AVIM.": GOTO 50096
50078  GOTO 50096
50080  GOSUB 50092: PRINT "IT IS NOT A VOICE PATTERNS": GOSUB 50100: PRINT
    "FILE.": GOTO 50096
50082 E = 1
50084  GOSUB 50092
50086  PRINT "FILE NOT FOUND."
50088  GOTO 50096
50090 C$ = "1": RETURN
50092  GOSUB 50102: PRINT "": HTAB 4: INVERSE : PRINT "!! ERROR";: NORMAL
    : PRINT " - ";: RETURN
50094 A$ = "VOICE PATTERN": RETURN
50096  INVERSE : PRINT "": HTAB 4: PRINT "   HIT ANY KEY TO CONTINUE";: NORMAL
    : GET B$: GOTO 50018
50098  PRINT "": PRINT "": HTAB 4: PRINT "";A$;"";" HAS BEEN LOADED": HTAB
    4: PRINT "TO THE AVIM.": GOTO 50096
50100  HTAB 4: PRINT "                ";: RETURN
50102  CALL  - 1059: CALL  - 1059: RETURN
50104 RT = (7 * 16) * 256: REM  SUBROUTINES ($7000)
50106 VF = (7 * 16 + 9) * 256: REM  VOCAB FILE
```

```
50108 RP = VF: REM REP. PATTERN
50110 IB = 512: REM  KEYBOARD INPUT BUFF.
50112 CB = (7 * 16 + 8) * 256: REM AVIM BUFF(7800)
50114 MW = VF:LNUM = MW + 1:FLAG = VF + 5
50116 GIVE = RT + 36:RECEIVE = RT + 39:UPLODE = RT + 27:INIT = RT + 9
50118 SN = CB:NB = CB + 2:OB = CB + 32:RB = SN + 160
50120 PN = 1
50122 UPLODE = RT + 27
50124  RETURN
50126 ,96
53000 D$ =  CHR$ (4)
53010  HOME
53050  PRINT "HIT ANY KEY WITHIN TEN SECONDS"
53060  PRINT "TO TRAIN THE DISC WITH YOUR VOICE"
53070  FOR J = 0 TO 10
53080 K = 10 - J
53090  PRINT K
53100 II = 15
53110  FOR I = 1 TO II
53120 KK = 5 ^ 5.6
53130  IF  PEEK (49168) > 127 GOTO 2
53140  NEXT I
53150  NEXT J
53160  HOME
60000  PRINT "THE VOICE INPUT MODULE IS NOW LOADED"
60010  PRINT "WITH THE VOICE PATTERNS FOR BANK STREET"
60020  PRINT "WRITER. PLEASE PLACE THE BANK STREET"
60030  PRINT "WRITER DISC IN THE DISC DRIVE AND PRESS"
60040  PRINT "'CONTROL - OPEN APPLE - RESET'"
```

APPENDIX A4

VOICE TUTOR

USER'S

MANUAL.

VOICE TUTOR

VERSION 1.0

USER'S MANUAL

System Requirements

Apple 11e with 64k RAM
one disk drive in Slot 6, Drive 1
and a Voice Input Module fitted in Slot 4.

Copyright

Acknowledgements

VIM is a registered trademark of Voice Machine Communications Inc.

CONTENTS

Chapter 1	**Introduction**	**01**
1.1	VIM and Voice Tutor	01
Chapter 2	**Using Voice Tutor**	**03**
2.1	The Display	03
2.1.1	The Sound Prompt	03
2.1.2	The Graphic	05
2.1.3	The Therapist's Permanent Prompt	05
2.1.4	The Therapist's Temporary Prompt	06
2.2	Using The Display	06
2.2.1	Background	06
2.2.2	Typical Exercise	07
Chapter 3	**The Therapist's Prompts**	**11**
3.1	The Display	11
3.2	Reject Threshold	12
3.2.1	Current Setting	13
3.2.2	Set Z	13
3.2.3	Auto U	14
3.2.4	Hi H	14
3.2.5	Lo O	15
3.3	Voice Patterns	15
3.3.1	Train T	16
3.3.2	Load L	18
3.3.3	Save S	19
3.4	Prompts	20
3.4.1	1,2,3,4,5	21
3.4.2	OK	22
3.4.3	A,B,C,D,E	23
3.4.4	Step P	24
3.4.5	Auto Step X	24
3.5	Miscellaneous	25
3.5.1	Help ?	26
3.5.2	Values V	26
3.5.3	Image G	27
3.5.4	Feedback F	29
3.5.5	Note R,W	29
3.5.6	Sound N	31

1. INTRODUCTION

1.1 VIM and Voice Tutor

The Voice Input Module (VIM), supplied by G.C.D. International, allows the user to perform keyboard functions using his voice. The VIM is a speaker dependent, isolated word recogniser. It will recognise up to 80 different isolated sounds spoken by a person who has trained the machine to recognise the particular sounds. When the VIM recognises a sound it enters a predetermined set of keystrokes into the Apple computer. The manual supplied with the VIM should be consulted for a full description of the VIM's options and for instructions on installation of the hardware.

Voice Tutor is a program which uses the VIM's sound recognition facility to assist a therapist in articulation training for children. Voice Tutor sets up a set of prompts and graphics which encourage the child to say particular sounds. The VIM recognises the sound and feeds information back to the Voice Tutor Program. Voice Tutor then modifies the graphics and prompts to show that the sound was entered either correctly or incorrectly. This reinforces the child's new speech habits and stimulates the child's interest allowing longer training sessions.

This manual assumes the user has installed the VIM in Slot 4, and the disc drive in Slot 6, Drive A of an Apple IIe with 64 kilobytes of memory. We also assume that the user is familiar with the correct method of inserting discs into the disc drive. We recommend that the user keeps the supplied discs as masters and makes copies for day to day use. In practice it is often a good idea to keep a separate disc for each child. Please refer to the manuals supplied with the VIM, disc drive, and Apple IIe computer for details on these procedures.

The VIM speech recogniser, used by Voice Tutor, is set up for a fixed frequency band, and matches sounds in a set way. Because of this, the performance of equipment varies between people. Trials with the equipment showed that its accuracy varied from 60% to 95% when assessing the same set of pairs of similar sounds spoken by different people. We can provide no reliable method for predicting whether the equipment will work with a given person.

2. USING VOICE TUTOR

2.1 The Display

With the power to the computer turned off, insert a copy of the Voice Tutor disc and close the disc drive door. Turn the power to the monitor and computer on. After one minute the disc drive will stop and a display will appear on the screen which will be similar to the display in Figure 1. The display is typical of all Voice Tutor displays except the help display which is discussed in section 3.5.1.

You may see that the display is made up of four parts:-

1. The Sound Prompt
2. The Graphic
3. The Therapist's Permanent Prompt
4. The Therapist's Temporary Prompt.

These four parts are discussed briefly below. Part 3, the Therapist's Permanent Prompt, is discussed in detail in Chapter 3 which explains all options available in Voice Tutor.

2.1.1 The Sound Prompt

The sound prompt is a white square located at the top of the screen which may be in any one of five different positions. The position of the sound prompt indicates which sound the computer expects to receive. Generally a piece of paper is prepared by the

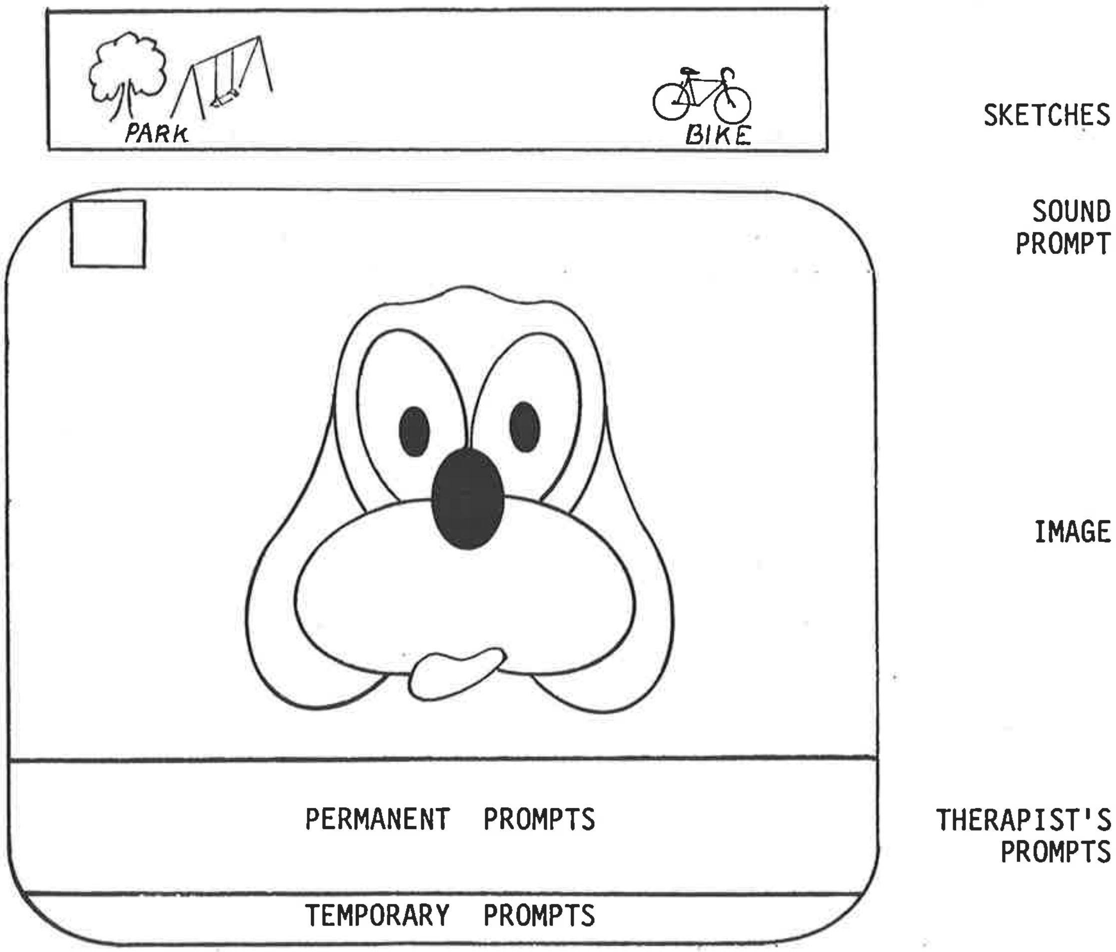


FIGURE 1. TYPICAL DISPLAY.

This figure shows the layout of a typical display on the monitor screen. A piece of paper is shown placed on top of the monitor screen. Two sketches are shown on the piece of paper representing the sounds "PARK" and "BIKE".

Therapist with sketches of objects which are meaningful to the client. The piece of paper is placed on the top of the monitor so that the sound prompt acts as a pointer to one of the five sketched objects. In practice, the client would say the sound associated with the sketch indicated by the sound prompt.

2.1.2 The Graphic

An image is located in the centre section of the screen. This image provides positive or negative feedback to the client depending on whether the recognised sound is the one that is expected. Each disc may have a maximum of ten images (however version 1.0 has only two images, a clown's face and a dog's head) which the Therapist may select to suit the client. (Image selection is discussed in Section 3.5.3). Take, for example, the clowns face. The client's correct pronunciation would produce a smile on the clown's face. A second correct pronunciation would add an extra part, for example a bow tie or hat, to the image. An incorrect pronunciation would produce a frown on the face. A second incorrect pronunciation would remove one of the parts from the image. (Note: negative changes may be suppressed as discussed in Section 3.5.4).

2.1.3 The Therapist's Permanent Prompt

The four lines at the bottom of the screen are filled with letters and numbers. The top three are the Therapist's Permanent Prompts which have three functions:-

1. To help the Therapist recall the available options.

2. To show the current status of various options.
3. To show the current value of variables which change in accordance with the client's performance.

The Therapist's permanent prompts cover all the options available and are discussed in detail in Chapter 3 of this manual.

2.1.4 The Therapist's Temporary Prompt.

Short messages are displayed on the bottom line of the screen from time to time. These messages prompt the Therapist through data entry procedures and assist the Therapist in keeping track of the various options. Generally messages shown on the bottom line are in response to the options shown in the Therapist's Permanent Prompt and are discussed in more detail in Chapter 3.

2.2 Using The Display

This section will provide an overview of the use of Voice Tutor. Various options will be mentioned in passing, refer to Chapter 3 for a detailed discussion on these.

2.2.1 Background

Voice Tutor is designed to be a game which is played with the voice. The player is successful, and is rewarded, whenever he pronounces a sound in a way that is similar to a particular sound stored in memory. The person setting up the game may use a number of options to set up the "rules" of the game. For example, the number of target sounds may be varied, the target sounds may alternate after each turn or may stay the same, the accuracy of pronunciation required by the computer may vary automatically, the player may be rewarded by a musical tune as well as a graphic, and several other variations.

However, there are several aspects of Voice Tutor which always stay the same and should be kept in mind when setting up the game. The following list summarises these aspects:-

1. The computer must have examples of the correct and incorrect pronunciations of the words which will be used. The computer may store a maximum of five sounds which will be called "library sounds" in this manual.
2. Each of the five library sounds corresponds with one of the five positions of the sound prompt.
3. When a sound is entered during the game, the computer compares the incoming sound with the five library sounds. If the incoming sound matches the library sound which is currently being indicated by the sound prompt then the pronunciation is taken to be correct. If the incoming sound is matched with one of the other library sounds then the pronunciation is in error.

2.2.2 Typical Exercise

The following is an example of a typical exercise. In this case a Speech Therapist sets up Voice Tutor for a client who tends to fail to pronounce the final consonant in words. Refer to Chapter 3 for a detailed explanation of the various options that are mentioned.

The Therapist selects two words which are examples of the words the client is inclined to mispronounce. The common mispronunciation of the word is also selected. In this case the words are "Park" and "Par", and "Bike" and "Buy". The Therapist loads and runs Voice Tutor so that a display similar to Figure 1 is shown in the monitor. The client is quite young and unable to read so the Therapist prepares symbols representing a park and a bike and places them over the sound prompt locations one and five (see Figure 1).

Only four sounds will be used, however the equipment has five library patterns. To prevent the fifth library pattern being confused with the other four, the Therapist moves the sound prompt to the centre of the screen (position 3) by pressing the "P" key. If the sound prompt refuses to move to any position then refer to Section 3.4.3. The Therapist then presses the "T" key and trains, with his own voice, a sound which is unlikely to be confused with the other four, for example the word "see". The client then trains the four sounds "Park, Par, Buy and Bike" into the positions 1,2,4 and 5. This is done by moving the sound prompt with the "P" key and then pressing the "T" key to train the particular sound.

Now that the library sounds are trained the Therapist sets up the game parameters by working through the options in the Therapist's Permanent Prompt. Working from left to right, a low reject threshold of 30 is entered with the "Z" key (Section 3.2.2) and the "U" in the word "Auto" is made lower case by pressing the "U" key (Section 3.2.3). This ensures that the reject threshold stays constant at a low value to let the client build up confidence by allowing a wide range of acceptable pronunciation. Moving to the centre of the Therapist's Permanent Prompt, the Therapist decides that the client should only be requested to say the words "Park" and "Bike". Therefore the second prompt should only stop at the first and fifth locations. To ensure this happens the Therapist presses the "A", "B", "C", "D" and "E" keys so that only "A" and "E" are shown as highlighted capitals. The Therapist decides that the client will be able to handle the sound prompt alternating between the two sounds so the "X" key is pressed until the "X" shown in the words "AUTO STEP X" is shown as a highlighted capital "X". Moving to the right end of the Therapist's Permanent Prompt, the Therapist decides that negative feedback should be initially suppressed to allow the client to gain confidence so the "F" key is pressed until the "F" in "FBACK" is shown as a lower case "f". Sound feedback is not required so the "N" key is pressed until the "N" in "SOUND" is shown as a lower case "n".

Voice Tutor is now set up and the client is invited to say the prompted sounds and gradually build up the graphic until a winning score is reached.

After some time the Therapist notes that the client is having trouble with the word "Park". The "X" key is pressed to allow the sound prompt to stay on the word "Park" so that the client can concentrate on this word. After several attempts the Therapist decides that the client's pronunciation is acceptable but the computer is not correctly recognising the word. The Therapist presses the "T" key and asks the client to retrain the word. After further use the therapist finds that the client is now easily entering the two sounds. The reject threshold is now increased using the "Z" key. This will mean that the pronunciation must be more precise or the computer will simply ignore the sound. The Therapist also presses the "F" key to introduce negative feedback so that errors in pronunciation will affect the graphic. The Therapist now notices that the client is mispronouncing the word "Park" in a different way, but the computer is accepting it as a correct pronunciation. The Therapist could overcome this problem by asking the client to retrain the "Park" and "Par" library sounds, or he could use the fifth library sound location (currently occupied by the Therapist's pronunciation of "See") to store an example of the new mispronunciation.

After a further period of use the time available for the exercise is complete, however the Therapist would like to continue the exercise on another day. In addition the client is close to finishing the image and winning the game. The Therapist therefore uses the "S" key to save the library sound pattern and the "V" key to save the current setup. The machine is turned off after doing this. On the following day, when the same disc is inserted and the computer turned on the game will be at the same place ready for continuation. The intention is that the client will, after some practice, and early monitoring, be able to carry out an articulation exercise with only partial assistance from the Therapist. This will allow the Therapist to attend to other tasks.

3. THE THERAPIST'S PROMPTS

3.1 The Display

```
REJ T zzz |TRAIN|12345 OK /|HELP?|VALUES
HIhhh AUTO|LOAD1|ABCDE STEP|IMAGEi|FBACK
LOooo SETZ|SAVEs|AUTO STEPX|NOTERW|SOUND
```

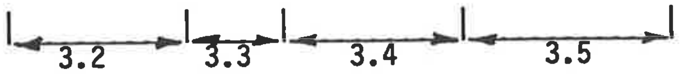


Figure 2. The Therapists Permanent Prompt Display

Keys associated with the characters shown in bold typing in Figure 2, and also the numbers 1 to 5, have some significance. Pressing the keys will activate or deactivate an option.

The Permanent Prompts may be considered in four logical segments. Figure 2 shows the segments and the section numbers in this manual which refer to the segments. The segments are:-

3.2 Reject Threshold

3.3 Voice Patterns

3.4 Prompts

3.5 Miscellaneous

These four segments explain in detail all the options available in Voice Tutor. Each option explanation is written so that it may be read by itself so that the explanations may be used for reference purposes. There is therefore considerable duplication in content and format. It is suggested that the reader scans Chapter 3 and returns to particular option descriptions when using the program to accomplish particular tasks.

In the following the phrase "at any time" is used in reference to activating options. This phrase means that the option may be selected at any time during operation of the main program. Attempts to enter a second option while in an option will be ignored by the computer or will cause the main option to be terminated. We are not aware of any instances where attempts to enter a second option will cause rejection of the overall program or loss of stored sound patterns or variables.

3.2 Reject Threshold

```
REJ T zzz |
HIhhh AUTO|
LOooo SETZ|
```

Figure 3. The Reject Threshold Prompt Display.

The reject threshold is a measure of the required accuracy of the subjects pronunciation. If the reject threshold is set very low (say at 20) then virtually any sound will be accepted by the computer and will be compared with the five sounds in memory. The computer will decide which of the five sounds is closest to the new sound and will assume the "best match" sound is the sound pronounced.

On the other hand, if the reject threshold is set very high (say at 120) then only pronunciations which are almost identical to one of the five stored sounds will be accepted by the computer.

The aim is to set the reject threshold so that poor pronunciations, background noises, etc are discarded by the computer. The Reject

Threshold Prompt displays the current value of the reject threshold and reminds the user how to change the reject threshold both manually and automatically. These options are detailed below.

3.2.1 Current Setting

The letters "zzz" in Figure 2 show where the current value of the reject threshold is displayed. The number will be between 10 and 125.

3.2.2 SET Z

The Z Key may be depressed at any time to manually set the reject threshold. When Z is depressed a message will appear at the bottom of the screen "INPUT REJECT THRESHOLD 3 DIGITS". Type in three digits between 10 and 125. Note however that numbers between 10 and 99 must be entered with a preceding 0, i.e. 010 and 099. The left arrow key may be used to erase digits. When the third digit is depressed the number will immediately be the new reject threshold and will be displayed in the zzz position on Figure 2. If an invalid entry is made then the user is returned to the main part of the program and may go on entering new sounds, or may try to enter a new reject threshold by depressing the Z key. A message will also be displayed at the bottom of the screen advising the user of the valid reject threshold range.

3.2.3 AUTO U

The U key may be depressed at any time to activate or deactivate the reject threshold automatic change option.

When this option is deactivated (shown by a lower case, un-highlighted, "u" in the word "AUTO" in Figure 3) then the reject threshold stays the same unless manually changed. However, if the option is activated (shown by an upper case, highlighted, "U" in the word "AUTO" in Figure 3) then the reject threshold increases every time a word is recognised as correct by the computer. A word recognised to be incorrect by the computer will cause the reject threshold to automatically decrease. This option might be used to make sound recognition progressively harder for the client to provide increasing challenge.

A message will appear when the "U' key is depressed to show whether the option has been activated or deactivated.

3.2.4 HI H

The automatically changing reject threshold described in 3.2.3 varies between a high and low limit. Depressing the "H" key will allow the Therapist to enter a three digit number between the low limit and 125 which will become the high limit. Note however that numbers between 10 and 99 must be entered with a preceding 0 i.e. 010 and 099. The left arrow key may be used to erase digits. When the third digit is depressed the number will immediately be the new high limit and will be displayed in the "hhh" position in Figure 2. If an invalid entry is made then the user is returned to the main part of the program and may go on entering new sounds, or may try to enter a new high limit by depressing the "H" key. A prompt will also be displayed at the bottom of the screen advising the user of the valid high limit range.

3.2.5 LO O

The automatically changing reject threshold described in 3.2.3 varies between a high and low limit. Depressing the "O" key will allow the Therapist to enter a three digit number between 10 and the high limit which will become the low limit. The notes and restrictions described in 3.2.4 apply here, however the low limit value will be displayed in the "ooo" part of Figure 2.

3.3 Voice Patterns

|TRAIN|

|LOAD1|

|SAVEs|

Figure 4. The Voice Patterns Prompt

Data representing five sounds are stored in the computers active memory. Incoming sounds or utterances are compared with these five library sounds. To be meaningful the incoming sounds must be similar to one or more of the library sounds. This is normally achieved by training the users voice patterns into the computer. Depressing the key "T", associated with the word "TRAIN" in the Voice Patterns Prompt, allows the user to train sounds into the active memory of the computer. The user may then use these library sounds, however they will be lost when the computer is turned off unless he uses the save option. The save option is activated by the key "S" associated with the word "SAVE" in the Voice Patterns Prompt. Once library sounds have been saved to disc they may be retrieved at a latter date with the Load option. The Load option is activated by the key "L" associated with the word "LOAD" in the Voice Patterns Prompt.

3.3.1 TRAIN T

Five sounds may be stored in the computers active memory. Each of these sounds is associated with one of the five positions of the sound prompt (the white square at the top of the screen). This association is established at the time of the sound training by the position of the sound prompt during the training. The user should be aware that training a sound into a position which is already occupied will destroy the old sound unless it has been saved to disc (see Section 3.3.2).

The first step in training a sound onto the computer is to move the sound prompt (the white square at the top of the screen) to the desired position. For example, if you want the word being trained to be associated with the sound prompt located at the far right of the screen, then press key "P" until the white square is at the far right of the screen, (key "P" is described in greater detail in Section 3.4.4).

Press the key "T" after locating the sound prompt in the desired position. "T" is associated with the word "TRAIN" in the voice patterns prompt (Figure 4) and is used to activate the training option. When "T" has been pressed the capital "T" in the voice patterns prompt at the bottom of the screen will change to a lower case "t" to show you are in the training mode, and a message will appear at the bottom of the screen stating "TRAINING - SAY THE PROMPTED WORD". The prompt is referring to the position of the sound prompt. If a piece of paper is arranged above the monitor screen with symbols representing the words being trained then the sound prompt will be pointing to the sound which is currently being trained. The word being trained should now be spoken into the microphone. Ensure the microphone is 30 to 60mm from the lips and pointed directly at the lips. The sound prompt will blink when a sound is detected. Three consecutive similar sounds are required before completing the training of a word. The computer stores the average of these three sounds. If the computer decides that the sound is similar to the last sound made then the sound prompt will blink once. If, however, the sound is not similar then the sound prompt will flicker. The user must simply continue repeating the sound until the computer detects three similar sounds in a row. Most people find that this is quite easy after a little practice.

The training routine is complete when the computer detects three consecutive similar sounds. The lower case "t" in the voice patterns prompt will change to upper case and a message will appear at the bottom of the screen stating "TRAINING FINISHED". You may now use the program in the normal way or train another sound by first moving the sound prompt.

3.3.2 LOAD L

If a set of sounds has already been saved to the disc (see Section 3.3.3) then they may be loaded into the computer's active memory using the load option. To use the load option press the "L" key. "L" is associated with the word "LOAD" in the voice patterns prompt (Figure 4). When "L" has been pressed the capital "L" in the voice patterns prompt at the bottom of the screen will change to a lower case "l" to show you are using the load option, and a message will appear at the bottom of the screen stating "# OF FILE = ? 1 DIGIT". This message is asking which of the ten sound pattern files (each containing five sounds) should be loaded into active memory. Pressing the return key at this stage will cancel the load option without making changes and will return the user to the main program. Any of the digits 0 to 9 may be entered as a legitimate file number. As soon as the digit key is pressed the disc drive will start and load the sound patterns into active memory. It should be noted that the patterns currently in active memory will be lost. The number of the file that was loaded is placed next to the word "LOAD" in the Voice Pattern Prompt (at the location of the lower case "l" in Figure 4) to remind the user of the number of the file currently in active memory.

When the computer is first turned on with the Voice Tutor disc in the disc drive the computer automatically loads one of the sound pattern files from the disc. The number of the voice pattern file which is automatically loaded depends on the setting nominated in the "values" option (see Section 3.5.2).

If the disc is not inserted in the disk drive properly, or if the disc has developed an error then an input/output (I/O) error will occur. The computer will reject the Voice Tutor Program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pushing "Control, open apple, reset" (see The Apple Computer Manual), or by turning the computer off and on.

3.3.3. SAVE S

If a set of sounds has already been trained into the computer's active memory (see Section 3.3.1) then the user may wish to save the sound patterns onto disc to prevent loss when the computer is turned off. The "SAVE" option is used to achieve this. Press the "S" key to use this option. "S" is associated with the word "SAVE" in the Voice Patterns Prompt (Figure 4). When "S" has been pressed the capital "S" in the Voice Patterns Prompt at the bottom of the screen will change to a lower case "s" to show you are using the save option, and a message will appear at the bottom of the screen stating "# OF FILE = ? 1 DIGIT". This prompt is asking you to nominate a number between 0 and 9 which will be the number of the file which will store the sound patterns currently in active memory. Pressing the return key at this stage will cancel the save option without making changes and will return the user to the main program. Any of the digits 0 to 9 may be entered as a legitimate file number. As soon as the digit key is pressed the disc drive will start and load the current active memory sound patterns onto the disc. It should be noted that the sound patterns which were on the disc under the nominated file number will be lost since the sound patterns from active memory will be written over them. The number of the file that was saved is placed next to the word "SAVE" in the Voice Pattern Prompt (at the location of the lower case "s" in Figure 4) to confirm the entered number for the user.

If the disc is not inserted in the disc drive properly, or if the disc has developed an error, or if the write protect sticker is on the disc (see the Computer Manual) then an input/output (I/O) error will occur. The computer will reject the voice tutor program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pushing "Control, open apple, restart" (see the Apple Computer Manual), or by turning the computer off and on.

3.4 Prompts

|12345 OK /|
|ABCDE STEP|
|AUTO STEPX|

Figure 5. The Prompt Options Display

The close relationship between the sound prompt, the five library voice patterns and the incoming sound has been discussed in detail in earlier sections, particularly Section 2.2. This section of the Therapist's Permanent Prompt assists the user in manipulating the sound prompt, and assists in assessing the subject's performance by indicating which sound was recognised. These options are presented in detail below.

3.4.1 1 2 3 **4** 5

The digits one to five appear in the top left hand corner of the "Prompt Options Display", Figure 5. These digits are used to show which sound was last recognised by the computer. For example, if a word is spoken into the microphone and the computer recognises it as the word associated with the left most position of the sound prompt then the digit 1 will be highlighted, digits 2 to 5 will be displayed normally. Similarly, if the next word spoken is recognised to be associated with the centre position of the sound prompt (i.e. third from the left) then the digit 3 will be highlighted.

This display is for use by the Therapist in understanding what the computer is receiving and doing. For example, if the subject says the word associated with the third sound prompt position but the number 5 is highlighted then the computer is saying that library pattern number 5 is a better match for the word just spoken than library pattern number 3. If this occurs a few times, and the Therapist feels that the word spoken was acceptable then consideration should be given to retraining either library pattern number 3 or 5.

The keys 1 to 5 have special significance. Pressing these keys has exactly the same effect as saying the corresponding word. For example, pressing the key 2 has exactly the same result as correctly saying the word which matches the second library pattern. This option might be used when demonstrating the program to the subject. For example, the subject's voice patterns may be in the library so that the Therapist would be unlikely to be able to say a word and have the program act in a predictable fashion. In this case the Therapist could demonstrate positive feedback by pressing the key corresponding to the present location of the sound prompt. i.e., If the sound prompt is in position 5 (to the far right of the screen) then pressing key 5 will cause:-

(a) The graphic to provide positive feedback (see Section 2.1.2).

(b) The number 5 will be highlighted on the Therapist's Permanent Prompt.

(c) A tick will appear next to the OK symbol (see 3.4.2).

(d) The sound prompt to move position, if that option is active (see Sections 3.4.3, 3.4.4 and 3.4.5).

3.4.2 OK

After each word is entered and accepted by the computer a diagonal line "/" or a cross "X" will appear next to the "OK" symbol in the "Prompt Option Display", Figure 5. These symbols show whether the last sound received matched with the library pattern which was indicated by the sound prompt. The diagonal line may be thought of as a tick for a correct match, and the cross an incorrect match. For example, if the sound prompt is located fourth from left then the computer expects to receive a sound which matches the fourth library pattern. If a sound is entered and accepted by the computer but does not match then an "X" will be displayed. If the sound does match with the fourth library pattern then a "/" will be displayed.

3.4.3 ABCDE

The sound prompt, in effect, requests the subject to say one of the five sounds. The Therapist may however wish that the subject should only say one or two of the five sounds. The Therapist uses keys A to E to define which words will be indicated by the sound prompt.

The letters A to E in the "Prompts Option Display" show which words will be indicated by the sound prompt. Letters highlighted indicate sound prompt positions which will be used. Letters shown normal in lower case indicate sound prompt positions which will not be used. The Therapist selects which positions will, or will not be indicated by the sound prompt by pressing the keys "A" to "E". A message will appear in the bottom line of the screen indicating whether the sound will or will not be prompted.

An example will illustrate this. If all the letters "A" to "E" are highlighted on the Therapist's Permanent Prompt then the sound prompt will appear in each of the five possible positions. If the "B" key is now pressed, the highlighting of the B in the Therapist's Permanent Prompt will disappear and be replaced by a lower case "b". A message will also appear at the bottom of the screen stating "THE "B" SOUND WILL NOT BE PROMPTED". If the sound prompt is now moved across the screen by repeatedly pressing the "P" key (see Section 3.4.4 for details) then you will see that the sound prompt does not stop at the number 2 position. So the second word from the left is not prompted. If the "B" key is now pressed again the capital "B" will reappear and will be highlighted. A message stating "THE "B" SOUND WILL BE PROMPTED" will appear at the bottom of the page. Repeated pressing of the "P" key will show that the sound prompt now stops at the number 2 position.

If A,B,C,D and E are pressed to prevent the sound prompt stopping at any position then the following message appears at the bottom of the screen "PRESS A,B,C,D OR E TO ACTIVATE A PROMPT". The sound prompt will not move, and the image will not change, until this request is carried out, however the computer will continue to accept and assess sounds.

3.4.4 STEP P

Pressing the "P" key forces the sound prompt to move to the right to the next active sound prompt position (see 3.4.3 to activate and deactivate the sound prompt positions). If there is only one active sound prompt position then the sound prompt will blink but stay in the same position. The step option is particularly useful when the auto stop option (see 3.4.5) has been changed to manual. For example, when the subject is required to repeat only one of the five sounds for some time. The Therapist may then step the sound prompt onto the next desired sound, when required, by pressing the "P" key.

3.4.5 AUTO STEP X

The Auto Step Option may be both activated and deactivated by pressing the "X" key. The Therapist's Permanent Prompt shows that the option is active by highlighting the "X" in the display. A lower case, normal "x" is displayed if the option is deactivated.

If the option is activated then the sound prompt automatically steps to the right to the next active position after each sound is accepted by the computer. (See 3.4.3 to activate and deactivate the sound prompt positions.) If the auto step option is deactivated then the sound prompt stays in the one position until manually moved with the "P" key (see 3.4.4).

The following example illustrates the use of this option. The Therapist has set up an exercise where the first and fifth sound prompt positions are active and the auto step option is active. The sound prompt will step from 1 to 5 and back to 1 as sounds are accepted by the computer. However the Therapist decides that the subject needs more practice on sound number 5. The auto step option is deactivated and the sound prompt moved to the number 5 position using the "P" key. The sound prompt will now blink after each sound is accepted but will not move away from the number 5 position.

3.5 MISCELLANEOUS

|HELP?|VALUES

|IMAGEi|FBACK

|NOTERW|SOUND

Figure 6. Miscellaneous Options Display

There are six additional options which assist the Therapist by:-

(a) providing an explanation of the option keys.

(b) allowing a note to be written to the disc.

(c) storing the current screen, options, and values so that the exercise may be completed on another day.

(d) allowing changes to the audiovisual feedback to the subject.

These options are discussed in detail below.

3.5.1 HELP ?

Pressing the "?" key will cause the computer to replace the current screen with a table which gives a short explanation of all the option keys. When finished using the table type any key to return to the main program. The computer will return to the same screen as before. You will notice that the disc drive will operate when using the help option. If the disc is not inserted in the disc drive properly, or if the disc has developed an error, or if the write protect sticker is on the disc (see the Computer Manual) then an input/output (I/O) error will occur. The computer will reject the Voice Tutor Program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pushing "Control, Open Apple, Reset" (see The Apple Computer Manual), or by turning the computer off and on.

3.5.2 VALUES **V**

During an exercise several options may be activated or deactivated, the reject value altered, and the image at a certain stage of development. It is possible to store this information on the disc by pressing the "V" key. This would allow you to turn the computer off and come back sometime later, re-insert the disc, turn the power on and return to exactly the same display. It should be noted however that pressing "V" does not store the current voice patterns, see Section 3.3.3 for details on the "Save" option which serves this function.

When "V" is pressed the disc drive operates to save the current information onto the disc. The following message appears at the bottom of the screen "CURRENT SETUP SAVED". When the disc drive stops you are returned to the main program and may proceed to alter the options if you wish. However, further changes will not be recorded on the disc unless you press "V" again.

If the disc is not inserted in the disc drive properly, or if the disc has developed an error, or if the write protect sticker is on the disc (see the Computer Manual) then an input/output (I/O) error will occur. The computer will reject the Voice Tutor Program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pressing "Control, Open Apple, Reset" (see The Apple Computer Manual), or by turning the computer off and on.

3.5.3 IMAGE **G**

Voice Tutor may have as many as ten images for visual feedback to the client, however version 1.0 is fitted with two. The Therapist changes the image by pressing the "G" key. The highlighted capital "G" is replaced by a normal, lower case "g" in the Therapist's Permanent Prompt to show the "Image" option has been activated. A message appears at the bottom of the screen stating " OF FILE = ? 1 DIGIT". The Therapist should enter a single digit from 0 to 9 to select an image. On version 1.0 the digit 0 will select a dog's head and digit 1 will select a clown's face. Pressing "Return" will return the user to the main program without making changes. After a digit key is pressed the disc drive will operate and a new image will appear on the screen. The entered digit will also appear to the right of the word "Image" on the Therapist's Permanent Prompt and the "G" in "Image" will become a highlighted capital "G".

This option may be used at any time, the image is automatically adjusted to reflect the progress of the client. For example, if the clown's face is being displayed with a smile and at the 10th stage of development, then the dog's head, if selected, would appear with a smile and at the 10th stage of development.

If the disc is not inserted in the disc drive properly, or if the disc has developed an error, then an input/output (I/O) error will occur. The computer will reject the Voice Tutor Program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pushing "Control, Open Apple, Reset" (see the Apple Computer Manual), or by turning the computer off and on.

3.5.4 FBACK F

The image will always be modified to produce positive feedback whenever a correct match is made between the incoming sound and the library pattern indicated by the sound prompt. Negative feedback will also be applied unless the "FBACK" option is deactivated.

The "FBACK" option is activated and deactivated by pressing the "F" key. For example, assume that the "FBACK" option is active. This will be apparent to the Therapist because the "F" in the word "FBACK" will be a highlighted capital. Also sound mismatches will cause the image to "frown" or to diminish providing a negative feedback. Press the "F" key to deactivate the "FBACK" option. A message will appear at the bottom of the screen stating "NO NEGATIVE FEEDBACK" and the F in "FBACK" will become a normal, lower case "f". Further sound mismatches will not cause a change in the image, however the sound prompt will flicker to indicate that a sound was received. To reactivate the negative feedback press the "F" key. A message stating "FEEDBACK ON" will appear at the bottom of the screen and the lower case "f" in "FBACK" will change to a highlighted capital "F".

3.5.5 NOTE**RW**

The "NOTERW" option allows the user to write a comment to the disc for latter reference. The note may be 40 characters long and may contain any alpha numeric data. It is suggested that the note might contain the client's name, or a file or cross reference number so that the voice patterns stored on the disc may be identified at a latter date.

To write a comment to the disc press the "W" key. A message will appear at the bottom of the screen stating "TYPE YOUR MESSAGE". You may now press any key to type your message which will appear on the bottom line of the screen. You may end your message and escape from the option by typing 41 characters or by pressing return. When this is done the disc drive will operate and the message will be saved to disc. You will be returned to the main menu. There are no editing facilities for this option, if an error is made simply press "return" and press "W" again after the disc drive stops. The disc can only store one message so that old messages will be lost when the new message is written.

To read the comment on the disc press the "R" key. The disc drive will operate and the message will appear at the bottom of the screen. You will be returned to the main program as soon as the disc drive stops.

If the disc is not inserted in the disc drive properly, or if the disc has developed an error, or if the write protect sticker is on the disc (see the Computer Manual) then an input/output (I/O) error will occur. The computer will reject the Voice Tutor Program and sound patterns currently in the active memory will be lost. In this case, rectify the error and restart from the beginning by pressing "Control, Open Apple, Reset" (see the Apple Computer Manual), or by turning the computer off and on.

3.5.6 SOUND N

As well as the visual feedback provided by the image, an aural feedback is available with the "SOUND" option. This option is activated by pressing the "N" key. When the "N" key is pressed a message appears at the bottom of the screen stating "sound on" and the n in "SOUND" in the Therapist's Permanent Prompt becomes a highlighted capital. Thereafter, whenever a correct match between the incoming sound and the library pattern indicated by the sound prompt occurs, two tones of increasing pitch are heard. If the "FBACK" option (Section 3.5.4) is active a single lower tone is heard when a mismatch of sounds is recognised by the computer. When sufficient correct matches are received by the computer the image becomes complete, a message stating "WINNING SCORE" is displayed and the computer plays a short tune.

When the "N" key is pressed again the "SOUND" option is deactivated. A message stating "SOUND OFF" is displayed at the bottom of the screen, and the highlighted capital "N" in "SOUND" is replaced by a lower case, normal "n".

APPENDIX A5

VOICE TUTOR

COMPUTER

PROGRAM.

```
]CATALOG

DISK VOLUME 254

 A 054 VOICE TUTOR
 B 013 GRAPHIC 0
*B 009 AVIMCOM
 A 002 HELLO
 A 003 AAB
 B 002 SLOT
 B 013 GRAPHIC 1
 A 009 AAC
 A 002 AAD
 B 013 GRAPHIC 2
 B 013 GRAPHIC 3
 B 005 VOICE TUTOR V
 A 006 AAF
 B 002 SPEAKER
 B 013 GRAPHIC 4
 B 013 GRAPHIC 5
 B 013 GRAPHIC 6
 B 013 GRAPHIC 7
 B 013 GRAPHIC 8
 B 013 GRAPHIC 9
 B 004 VOICE TUTOR VP 0
 B 004 VOICE TUTOR VP 1
 B 004 VOICE TUTOR VP 2
 B 004 VOICE TUTOR VP 3
 B 004 VOICE TUTOR VP 4
 B 004 VOICE TUTOR VP 5
 B 004 VOICE TUTOR VP 6
 A 003 AAH
 B 004 VOICE TUTOR VP 7
 B 004 VOICE TUTOR VP 8
 B 004 VOICE TUTOR VP 9
 B 005 TITLE
 A 004 IMAGE COMPILER
 B 013 GRAPHIC A
 B 004 PROMPTS
 B 002 TOGGLES
 B 002 PERM PROMPTA
 B 002 PERM PROMPTB
 B 002 PERM PROMPTC
 B 005 HELP
 B 005 TEMP
 B 002 MESSAGE

]
```

```
]LOAD HELLO
]LIST

10  HIMEM: 24999
20 D$ =  CHR$ (4)
30  GR
40  PRINT D$;"BLOAD TITLE"
50  PRINT D$;"RUN VOICE TUTOR"
```

```
LOAD VOICE TUTOR
]LIST

1  HIMEM: 24999
2 D$ =  CHR$ (4)
10  GOSUB 30000
20  PRINT D$;"BLOAD AVIMCOM,A";RT
30  GOSUB 40000
40  GOTO 50000
100  REM
110  POKE OB,24: POKE NB,1: CALL GIVECMD: POKE N2,3: CALL RECEIVE:ST =  FN
    LH(RB): RETURN
210  POKE OB,16: POKE OB + 1, PEEK (RP): POKE OB + 2, PEEK (RP + 1): POKE
    NB,3: CALL DNVPVC
220  POKE OB,4: POKE OB + 1, PEEK (RJ): POKE NB,2: CALL GIVECMD: RETURN
300  REM
310  GOSUB 100: IF ST = 0 OR ST = 1 THEN  POP : GOTO 20300
320  POKE OB,25: POKE OB + 1,0: POKE OB + 2,0: POKE CB + 2,3: POKE N2,6
330  CALL UPLODE: POKE VF + 7,85: POKE VF + 8,170: RETURN
400  REM
405  PRINT : HTAB 4: INPUT "TYPE REJECT THRESHOLD(0 TO 128)";A$: ON  NOT
     LEN (A$) GOTO 420:I =  VAL (A$): IF I < 0 OR I > 128 THEN  PRINT BE$
    : GOTO 400
410  POKE OB,4: POKE OB + 1,I: POKE NB,2: CALL GIVECMD
420  RETURN
500  REM
505  POKE OB,5: POKE NB,1: CALL GIVECMD: POKE N2,1: CALL RECEIVE
510  PRINT : HTAB 4: PRINT "CURRENT REJECT THRESHOLD="; PEEK (RB): GOTO 9
    260
600  REM  LOAD VOICE TUTOR VP
602 T1 = 170:T2 = 85:V$ = "R":ZZ = 7
605 A$ = "VOICE TUTOR VP"
620  PRINT D$;"BLOAD ";A$;",A";VF:VF$ = A$
630  GOTO 3200
700  REM
705 A$ = "VOICE TUTOR VP " + B$
720  PRINT D$;"BSAVE ";A$;",A";A;",L";L
730  RETURN
802  GOSUB 100: IF ST <  > 257 GOTO 20600
805  HOME : VTAB 10: GOSUB 8600: IF DV% < 0 THEN  RETURN
810  IF DV% AND  NOT PN THEN  RETURN
815  GOSUB 300
820  POKE VF + 6,DV%: IF DV% = 1 THEN  PRINT D$;"PR#";PN
825  GOSUB 9700:WN = 1
830  POKE OB,8: POKE OB + 1,1: POKE OB + 2,252: POKE OB + 3, PEEK (MW): POKE
    OB + 4,255: POKE NB,5: CALL GIVECMD
835  POKE N2,4: CALL TESTVC:I =  PEEK (RB): IF I = 251 THEN  RETURN
840 WN% =  PEEK (RB): IF WN% > 251 THEN  PRINT "***  REJECTED  ***": GOTO
    880
845  POKE CW,WN%: GOSUB 9620: CALL PWRD
850 WN% =  PEEK (RB + 3): POKE CW,WN%: HTAB 19: CALL PWRD
855 I =  PEEK (RB + 1): HTAB 36: PRINT I
860 GS =  PEEK (RB + 2): IF GS < MS THEN MS = GS
865 TS = TS + GS:WC = WC + 1
870  IF I < MD THEN MD = I
875 TD = TD + I
880 WN = WN + 1
885 MM% =  PEEK (MW): IF WN > MM% THEN WN = 1
```

```
890  GOTO 835
902  GOSUB 100: IF ST <  > 257 GOTO 20600
905  HOME : VTAB 10: GOSUB 8600: IF DV% < 0 THEN  RETURN
910  IF DV% AND  NOT PN THEN  RETURN
920  POKE VF + 6,DV%: GOSUB 300
930  IF DV% = 1 THEN  PRINT D$;"PR#";PN
940  GOSUB 9720: POKE OB,26: POKE OB + 1,5: POKE NB,2: CALL GIVECMD
950 DI = 3: POKE N2,4
960  CALL RECEIVE:I =  PEEK (RB): IF I = 255 THEN  PRINT : PRINT : HTAB 1
    0: PRINT "END OF TEST": PRINT D$;"PR#0": GOTO 9260
970 WN% =  PEEK (RB): POKE CW,WN%: GOSUB 9620: CALL PWRD:WN% =  PEEK (RB +
    3): POKE CW,WN%: HTAB 19: CALL PWRD: HTAB 36: PRINT  PEEK (RB + 1);: IF
     PEEK (RB + 1) < DI THEN  HTAB 39: PRINT "X";
980  PRINT : GOTO 960
3200  GOSUB 100: IF ST <  > 257 THEN  GOSUB 20600: GOTO 3000
3220  POKE OB,21: POKE NB,1: CALL GIVECMD
3230  POKE N2,0: CALL RECEIVE
3240  IF  PEEK (RB) <  > 67 THEN  GOSUB 20200: GOTO 3000
3340  RETURN
4000  REM
4005  GOSUB 300: ON ST <  > 257 GOTO 20600: POKE OB,18
4010  GOTO 4215
4215 I = I + 1
4220 J =  PEEK (MW): IF I > J GOTO 20400
4230  POKE OB + 1,3: POKE OB + 2,I: POKE OB + 3,I: POKE OB + 4,255: POKE
    NB,5
4240  POKE N2,0: CALL GIVECMD: CALL RECEIVE:A$ =  CHR$ ( PEEK (RB)): IF A
    $ = "E" GOTO 20500
4250  GOSUB 4820:XT = 3: GOSUB 7570: RETURN
4702  HOME : VTAB 5: PRINT : HTAB 5: IF TC = 19 THEN  PRINT "ADDITIONAL "
    ;
4704  IF TC = 18 THEN  PRINT "NEW ";
4705  PRINT "TRAINING": GOSUB 9230: IF B$ =  CHR$ (13) THEN  RETURN
4710  POKE OB,TC: POKE OB + 1,1: POKE OB + 2,255: POKE NB,3: CALL GIVECMD
    : GOSUB 4800
4720  IF TC = 18 THEN I = 106: GOSUB 410
4730  GOTO 9260
4800  REM
4810  POKE N2,0: CALL RECEIVE:A$ =  CHR$ ( PEEK (RB))
4820  CALL 64680: IF A$ = "C" THEN  GOSUB 51700: GOTO 4800
4830  IF A$ = "E" THEN  GOSUB 51700: GOSUB 51700: GOSUB 51700: GOSUB 5170
    0: GOTO 4800
4840  GOSUB 7570: RETURN
4900  REM  SAVE VP & VC
4905  GOSUB 100: ON ST <  > 257 GOTO 20600:ZZ = 5
4910 VF$ = "": POKE OB,20: POKE OB + 1,0: POKE OB + 2,0: POKE NB,3
4920  CALL LODEVPVC:V$ = "R": POKE VF + 7,170: POKE VF + 8,85
4930  POKE OB,5: POKE NB,1: CALL GIVECMD: POKE N2,1: CALL RECEIVE: POKE R
    J, PEEK (RB)
4940 A = RP:L =  FN HL(RP) + 32
4950 E = 0: GOSUB 700
4960  GOSUB 3200
4970  GOTO 51800
7570  FOR XC = 1 TO XT: CALL 64680: NEXT XC:XC =  FRE (0): RETURN
8210  GOSUB 8800: POKE CW,WN%: IF EF% = 1 THEN  RETURN
30000  REM
30005 RT = (7 * 16) * 256
30010 CLRMEM = RT
```

```
30015 S2ETCW = RT + 3
30020 UPLODE = RT + 6
30025 INIT = RT + 9
30030 NEVOCAB = RT + 12
30035 VF = (7 * 16 + 9) * 256
30040 PRNTVC = RT + 15
30045 ADDVC = RT + 18
30050 DLTECW = RT + 21
30055 RP = VF
30060 ISERTCW = RT + 24
30065 DNVPVC = RT + 27
30070 LODEVPVC = RT + 30
30075 DOWNLODE = RT + 33
30080 IB = 512
30085 GIVECMD = RT + 36
30090 RECEIVE = RT + 39
30095 PV = RT + 42
30100 MI = RT + 45
30105 CB = (7 * 16 + 8) * 256
30110 OKAY = CB + 1
30115 TESTVC = RT + 48
30120 N2 = CB + 3
30125 MW = VF:CW = MW + 1:EDXIT = VF + 5:DEVICE = VF + 6
30130 CF = VF + 10
30135 RJ = VF + 9
30140 PWRD = RT + 51
30145 COVERT = RT + 57
30150 CHANGECW = RT + 54
30155 SN = CB:NB = CB + 2:OB = CB + 32:RB = CB + 160
30160 MEMSIZE = (80 * 67) + 1600
30165 D$ =  CHR$ (4):BE$ =  CHR$ (7)
30170 S =  PEEK (43626):DR =  PEEK (43624)
30175  DEF  FN LH(I) =  PEEK (I) +  PEEK (I + 1) * 256: DEF  FN MF(I) = M
    EMSIZE -  FN LH(CF) - (72 *  PEEK (MW)) - I - 130: DEF  FN HL(I) =  PEEK
    (I) * 256 +  PEEK (I + 1)
30180 PN = 0
30185  RETURN
40000  PRINT D$;"BLOAD SLOT,A";SN
40010  CALL INIT:I =  PEEK (OKAY)
40020  IF I = 0 THEN  RETURN
40030  GOSUB 9190: VTAB 10: HTAB 4: PRINT "IF YOUR AVIM IS NOT "
40040  HTAB 4: PRINT "PLUGGED-IN AT APPLE I/O SLOT ";( PEEK (SN) - 128) /
    16
40050  PRINT : HTAB 4: PRINT "PLEASE TYPE THE CURRENT SLOT": HTAB 4: PRINT
    "NUMBER";
40060  INPUT A$:I =  VAL (A$): IF I < 1 OR I > 7 GOTO 40050
40070 I = 128 + I * 16: POKE CB,I: CALL INIT: IF  PEEK (OKAY) GOTO 40030
40080  PRINT D$;"BSAVE SLOT,A";SN;",L8"
40090  RETURN
45000  HOME : VTAB 4
45005  HTAB 7: PRINT "SET/READ PARAMETERS"
45010  HTAB 7: PRINT "*******************": PRINT ""
45015  HTAB 4: PRINT " #        MENU"
45020  HTAB 4: PRINT "--    --------------------"
45025  HTAB 4: PRINT "1 - SET REJECT THRESHOLD": PRINT
45030  HTAB 4: PRINT "2 - READ REJECT THRESHOLD": PRINT
45035  HTAB 4: PRINT "3 - SET WORD PARAMETERS": PRINT
45040  HTAB 4: PRINT "4 - READ WORD PARAMETERS": PRINT
45045  HTAB 4: PRINT "5 - EXIT"
```

```
45050  HTAB 4: PRINT "--  --------------------"
45055  PRINT : GOSUB 9090
45060  ON MN GOSUB 400,500,45300,45500
45065  IF MN = 5 THEN  RETURN
45070  GOTO 45000
45300  HTAB 4: INPUT "TYPE T1=";I: POKE OB,9: POKE OB + 1,I: POKE NB,7
45310  HTAB 4: INPUT "TYPE T2=";I: POKE OB + 2,I
45320  HTAB 4: INPUT "TYPE ETHL=";I: POKE OB + 3,I
45330  HTAB 4: INPUT "TYPE MIN. SAMPLE NUMBER=";I: POKE OB + 4,I: HTAB 4:
     INPUT "TYPE NOISE SAMPLE NUMBER=";I: POKE OB + 5,I
45340  HTAB 4: INPUT "TYPE NOISE THRESHOLD=";I: POKE OB + 6,I: CALL GIVEC
    MD: RETURN
45500  POKE OB,10: POKE NB,1: CALL GIVECMD: POKE N2,6: CALL RECEIVE: HOME
    : VTAB 6
45510  HTAB 4: PRINT "T1= "; PEEK (RB): PRINT : HTAB 4: PRINT "T2= "; PEEK
    (RB + 1): PRINT
45520  HTAB 4: PRINT "ETHL= "; PEEK (RB + 2): PRINT : HTAB 4: PRINT "MIN.
     SAMPLE= "; PEEK (RB + 3)
45530  PRINT : HTAB 4: PRINT "NOISE SAMPLE NUMBER= "; PEEK (RB + 4): PRINT
    : HTAB 4: PRINT "NOISE THRESHOLD= "; PEEK (RB + 5): GOTO 9260
49900  TEXT : HOME
50000  REM
50007 D$ =  CHR$ (4)
50010  DIM TG(28),IM(30)
50110  GOSUB 50200
50113 B$ =  STR$ (TG(22))
50115  GR
50120  GOSUB 50400
50122  GOSUB 60400
50123  PRINT D$;"BLOAD SPEAKER"
50125  GOSUB 59100: GOSUB 59600: GOSUB 59900
50127  POKE 1759,(TG(22) + 176)
50128  POKE 1777,(TG(21) + 176)
50130  GOSUB 50500
50140  GOSUB 51000
50180 B$ =  STR$ (TG(22))
50190  GOTO 53120
50200  REM
50210  PRINT D$;"BLOAD TOGGLES"
50220  FOR I = 0 TO 26
50230 TG(I) =  PEEK (28668 - I)
50240  NEXT I
50250  RETURN
50305  FOR I = 0 TO 26
50310  POKE (28668 - I),TG(I)
50320  NEXT I
50330  PRINT D$;"BSAVE TOGGLES,A28641,L28"
50340  RETURN
50400  REM
50410  PRINT D$;"BLOAD PERM PROMPTA"
50420  PRINT D$;"BLOAD PERM PROMPTB"
50430  PRINT D$;"BLOAD PERM PROMPTC"
50440  RETURN
50500  REM
50510  PRINT D$;"BLOAD PROMPTS"
50520  RETURN
51000  REM
51020  GOSUB 60000
51045  IF TG(14) < TG(20) THEN  GOTO 51055
```

```
51050 TG(14) = TG(20)
51055  IF TG(14) < 6 THEN TG(14) = 6
51060  FOR QZ = 1 TO TG(14)
51065  GOSUB 60200
51070  NEXT QZ
51445  RETURN
51450  REM
51455 I = TG(8)
51460  POKE OB,4: POKE OB + 1,I
51465  POKE NB,2: CALL GIVECMD
51475  RETURN
51500  REM
51510  IF TG(18) > 0 GOTO 51590
51520  REM
51530 TG(17) = 0
51540 TG(16) =  INT (TG(16))
51555  IF TG(17) > 6 GOTO 51600
51560 TG(16) = TG(16) + 1
51570  IF TG(16) > 5 THEN TG(16) = 0: GOTO 51560
51580  IF TG(TG(16)) = 0 THEN TG(17) = TG(17) + 1: GOTO 51555
51590  GOSUB 51700: GOTO 51800
51600  REM
51630  GOSUB 58100
51640  GOTO 51822
51700  REM
51710 TG(17) = 9 * (TG(16) - 1)
51720  FOR JJ = 0 TO 2
51730  COLOR= 0
51740  HLIN 0,39 AT JJ
51750  COLOR= 15
51760  HLIN TG(17),TG(17) + 3 AT JJ
51770  NEXT JJ
51780  RETURN
51790  GOSUB 60400
51800  REM
51820  GOSUB 51700
51822  REM
51825  GOSUB 51830: GOTO 51875
51830  REM
51835  POKE 49168,22
51840  IF  PEEK (49167) > 127 GOTO 51860
51850  GOTO 51840
51860 JJ =  PEEK (49168)
51870  IF JJ > 127 GOTO 51860
51872  RETURN
51875  REM
51880  IF JJ > 48.5 AND JJ < 53.5 GOTO 55000
51890  IF JJ > 64.5 AND JJ < 69.5 GOTO 52000
51895  IF JJ > 96.5 AND JJ < 101.5 GOTO 52000
51900  IF JJ > 90 THEN JJ = JJ - 32
51913  IF JJ = 63 GOTO 56000
51915  IF JJ = 70 GOTO 56300
51920  IF JJ = 71 GOTO 53700
51925  IF JJ = 72 GOTO 57200
51935  IF JJ = 76 GOTO 53100
51938  IF JJ = 78 GOTO 52100
51939  IF JJ = 79 GOTO 57100
51940  IF JJ = 80 GOTO 51520
51945  IF JJ = 82 GOTO 56250
51950  IF JJ = 83 GOTO 53500
```

```
51955  IF JJ = 84 GOTO 55800
51960  IF JJ = 85 GOTO 57000
51962  IF JJ = 86 GOTO 56500
51965  IF JJ = 87 GOTO 56100
51970  IF JJ = 88 GOTO 52300
51975  IF JJ = 90 GOTO 52410
51980  IF JJ = 8 GOTO 51800
51998  GOSUB 58200: GOTO 51800
52000  REM
52010  IF JJ > 70 GOTO 52040
52020 JJ = JJ - 64
52030  GOTO 52050
52040 JJ = JJ - 96
52050 JJ =  INT (JJ)
52060  IF JJ < 1 OR JJ > 5 GOTO 51800
52070 TG(17) = TG(JJ)
52080  IF TG(17) < 0.5 THEN TG(JJ) = 1: GOSUB 58300: GOSUB 60400: GOTO 51
   500
52090  IF TG(17) > 0.5 THEN TG(JJ) = 0
52095  GOSUB 58400: GOSUB 60400: GOTO 51500
52100  REM  SOUND TOGGLE
52110  GOSUB 58000
52120  POKE 2000,211: POKE 2001,207
52130  POKE 2002,213: POKE 2004,196: POKE 2006,207: POKE 2003,206
52140  IF TG(23) < 0.5 THEN TG(23) = 1: POKE 2007,206: GOTO 51790
52150 TG(23) = 0: POKE 2007,198: POKE 2008,198
52160  GOTO 51790
52300  REM
52310 JJ = TG(18)
52320  IF JJ < 0.5 THEN TG(18) = 1: GOSUB 58500: GOSUB 60400: GOTO 51500
52330 TG(18) = 0: GOSUB 58600: GOSUB 60400: GOTO 51500
52400  REM
52410  GOSUB 58000
52412 JM = 21:JL = 28109
52414  GOSUB 58050
52420 JI(4) = 3
52430  GOSUB 58700
52440  IF JI(5) < 10 OR JI(5) > 125 THEN  GOSUB 57900: GOTO 51800
52450 TG(8) =  INT (JI(5))
52460  GOSUB 51450
52470  GOSUB 59100
52495  GOTO 51790
53000  REM  GET # OF FILE
53010  GOSUB 58900
53020 JI(4) = 1
53030  GOSUB 58700
53040  IF JI(5) < 0 OR JI(5) > 9 THEN  GOTO 58200
53045 TG(22) = JI(5)
53050 B$ =  STR$ (JI(5))
53060  POKE 2012,(JI(5) + 176)
53070  RETURN
53100  REM  LOAD VOICE PATTERNS
53105  POKE 1755,236
53110  GOSUB 53000
53112  POKE 1755,12: POKE 1759,(JI(5) + 176)
53120  CALL INIT
53130 T1 = 170:T2 = 85:V$ = "R"
53140 I =  PEEK (512): IF E = 1 THEN  PRINT "ERROR 53140"
```

```
53150  PRINT D$;"BLOAD VOICE TUTOR VP ";B$;",A";VF
53160 TY =  PEEK (VF + 7): IF TY <  > T1 THEN  PRINT "ERROR 53160"
53170 TY =  PEEK (VF + 8): IF TY <  > T2 THEN  PRINT "ERROR 53170"
53180  POKE OB,16: POKE OB + 1,( PEEK (RP))
53190  POKE OB + 2,( PEEK (RP + 1)): POKE NB,3
53200  CALL (RT + 27)
53210 I = 106
53220  POKE OB,4: POKE OB + 1,I: POKE NB,2
53230  CALL GIVECMD
53340  GOSUB 3220
53350  GOTO 51590
53500  REM  SAVE VOICE PATTS
53505  POKE 1883,243
53510  GOSUB 53000: REM   GET FILE #
53515  POKE 1883,19: POKE 1887,(JI(5) + 176)
53530  GOTO 4900
53700  REM  LOAD GRAPHIC
53705  POKE 1775,231
53710  GOSUB 53000: REM   GET FILE #
53720 TG(21) = JI(5)
53730  GOSUB 51000
53735  POKE 1775,7: POKE 1777,(JI(5) + 176)
53740  GOTO 51800
55000  REM
55005 JJ = JJ - 48
55006  POKE (1632 + TG(24)),(176 + TG(24))
55007 TG(24) = JJ
55008  POKE (1632 + TG(24)),(48 + TG(24))
55010  IF JJ = TG(16) GOTO 55200
55025  POKE 1642,248
55030  IF TG(19) = 1 THEN TG(13) = 0: GOTO 55300
55032  IF TG(23) = 0 GOTO 55040
55034  POKE 864,2: POKE 865,(TG(14) * 3 + 100): CALL 866
55040 TG(12) = 0
55048 QZ = 1
55052  GOSUB 60200
55060  IF TG(13) = 0 AND TG(14) > 3 THEN TG(14) = TG(14) - 1
55070  IF TG(13) <  > 0 THEN TG(13) = 0: GOTO 55090
55074 QZ = TG(14) + 1
55078  GOSUB 60300
55080 QZ = TG(14)
55084  GOSUB 60200
55090  GOTO 55300
55200  REM
55205  POKE 1642,175
55210 TG(12) = 0
55216 QZ = 2
55220  GOSUB 60200
55222  IF TG(23) = 0 GOTO 55230
55224  POKE 864,2: POKE 865,(TG(14) * 3 + 100): CALL 866
55226  FOR I = 1 TO 5: NEXT I
55228  POKE 864,4: POKE 865,(TG(14) * 3 + 150): CALL 866
55230  IF TG(13) = 1 THEN TG(14) = TG(14) + 1
55235  IF TG(14) > TG(20) THEN  GOTO 55300
55240 TG(13) = 1
55246 QZ = TG(14)
55250  GOSUB 60200
55300  REM
55310  IF TG(14) > TG(20) THEN TG(14) = TG(20): GOSUB 59000: GOTO 51500
```

```
55320  IF TG(6) < 0.5 GOTO 51500
55330  IF TG(13) < 0.5 GOTO 55400
55340  IF TG(9) < 10 OR TG(9) > 125 THEN TG(9) = 125
55350 TG(8) = TG(8) + 10 * ( COS (TG(8) * 1.57079 / TG(9)))
55360 TG(8) =  INT (TG(8))
55370  IF TG(8) > TG(9) THEN TG(8) = TG(9)
55380  GOSUB 59100: GOSUB 51450
55390  GOTO 51500
55400  IF TG(9) < 10 OR TG(9) > 125 THEN TG(9) = 125
55410 TG(8) = TG(8) - 10 * ( COS (TG(8) * 1.57079 / TG(9)))
55420  IF TG(8) < 10 THEN TG(8) = 10
55430 TG(8) =  INT (TG(8))
55440  IF TG(8) < TG(7) THEN TG(8) = TG(7)
55450  GOSUB 59100: GOSUB 51450
55460  GOTO 51500
55800  REM
55802  POKE 1627,244
55805 I = TG(16)
55810  GOSUB 58000
55820 JM = 32:JL = 28160
55830  GOSUB 58050
55860  GOSUB 4000
55910  GOSUB 58000
55920 JM = 17:JL = 28127
55930  GOSUB 58050
55935  POKE 1627,20
55950  GOSUB 3200
55960  GOTO 51800
56000  REM
56010  PRINT D$;"BSAVE TEMP,A1024,L1016"
56015  TEXT : HOME
56020  PRINT D$;"BLOAD HELP"
56030  GOSUB 51830
56035  GR
56040  PRINT D$;"BLOAD TEMP"
56050  GOTO 51800
56100  REM
56110  GOSUB 58000
56120 JM = 16:JL = 28037
56130  GOSUB 58050
56140  FOR I = 0 TO 38
56150  GOSUB 51830
56160  IF JJ = 13 GOTO 56220
56170  POKE (2000 + I),(JJ + 128)
56180  POKE (2001 + I),223
56190  NEXT I
56200  GOSUB 51830
56210  POKE 2039,(JJ + 128)
56220  PRINT D$;"BSAVE MESSAGE,A2000,L40"
56230  GOTO 51800
56250  REM
56260  PRINT D$;"BLOAD MESSAGE"
56270  GOTO 51800
56300  REM
56310  IF TG(19) > 0.5 GOTO 56400
56330 TG(19) = 1
56340  GOSUB 58000
56350 JM = 19:JL = 28087
56360  GOSUB 58050
```

```
56370  GOTO 51790
56400  REM
56410 TG(19) = 0
56420  GOSUB 58000
56430 JM = 10:JL = 28067
56440  GOSUB 58050
56450  GOTO 51790
56500  REM
56510  FOR I = 0 TO 26
56520  POKE (28668 - I),TG(I)
56530  NEXT I
56540  PRINT D$;"BSAVE TOGGLES,A28641,L28"
56550  GOSUB 58000
56560 JM = 19:JL = 28020
56570  GOSUB 58050
56580  GOTO 51800
56600  REM
56610  GOSUB 58200
56620  GOSUB 58000
56630 JM = 18:JL = 28055
56640  GOSUB 58050
56650  GOTO 51800
57000  REM
57010  GOSUB 58000
57020  IF TG(6) < 0.5 THEN TG(6) = 1: GOTO 59200
57030 TG(6) = 0
57040  GOTO 59300
57100  REM
57110  GOSUB 58000
57120  GOSUB 59400
57130 JI(4) = 3
57140  GOSUB 58700
57150  IF JI(5) < 10 OR JI(5) > TG(9) THEN  GOSUB 59500: GOTO 51800
57160 TG(7) =  INT (JI(5))
57165  GOSUB 59600
57170  IF TG(8) < TG(7) THEN TG(8) = TG(7): GOTO 52460
57180  GOTO 51800
57200  REM
57210  GOSUB 58000
57220  GOSUB 59700
57230 JI(4) = 3
57240  GOSUB 58700
57250  IF JI(5) > 125 OR JI(5) < TG(7) THEN  GOSUB 59800: GOTO 51800
57260 TG(9) =  INT (JI(5))
57265  GOSUB 59900
57270  GOTO 51800
57900  REM
57910 JM = 39:JL = 28200
57920  GOSUB 58050
57930  RETURN
58000  REM
58010  FOR JK = 0 TO 39
58020  POKE 2000 + JK,160
58030  NEXT JK
58040  RETURN
58050  REM
58060  FOR JK = 0 TO JM
58070  POKE 2000 + JK,( PEEK (JL - JK))
```

```
58080  NEXT JK
58090  RETURN
58100  REM
58110  GOSUB 58000
58120 JM = 39:JL = 28640
58130  GOSUB 58050
58140  RETURN
58200  REM
58210  GOSUB 58000
58230 JM = 39:JL = 28600
58240  GOSUB 58050
58250  GOTO 51822
58300  REM
58310  GOSUB 58000
58320 JM = 39:JL = 28560
58330  GOSUB 58050
58340  POKE 2005,128 + JJ
58350  RETURN
58400  REM
58410  GOSUB 58310
58420  POKE 2019,206: POKE 2020,207
58430  POKE 2021,212: POKE 2022,160
58440  RETURN
58500  REM
58510 JM = 39:JL = 28520
58520  GOSUB 58050
58525  POKE 2006,16
58530  RETURN
58600  REM
58610  GOSUB 58000
58620 JM = 39:JL = 28480
58630  GOSUB 58050
58640  RETURN
58700  REM
58705  POKE 2028,JI(4) + 176
58710  POKE 2030,132: POKE 2031,137: POKE 2032,135
58715  POKE 2033,137: POKE 2034,148
58717  IF JI(4) < 2 GOTO 58720
58719  POKE 2035,147
58720  FOR JJ = 0 TO 3
58725 JI(JJ) = 0
58730  NEXT JJ
58735 JI(0) = 3 - JI(4)
58740 JI(0) = JI(0) + 1
58745  POKE 2036 + JI(0),32
58750  REM
58752  GOSUB 51830
58755  IF JJ = 8 GOTO 58800
58760  IF JJ = 13 GOTO 58825
58765  IF JJ < 48 OR JJ > 57 GOTO 58750
58775 JI(JI(0)) = JJ - 48
58777 JJ = JJ + 128
58780  POKE 2036 + JI(0),JJ
58785  IF JI(0) < JI(4) GOTO 58740
58790 JI(5) = JI(1) * 100 + JI(2) * 10 + JI(3)
58795  RETURN
58800  IF JI(0) = 1 GOTO 58825
58810  POKE 2036 + JI(0),160
```

```
58815 JI(0) = JI(0) - 1
58820  GOTO 58745
58825  REM
58830  GOSUB 58000
58835  POKE 2000,142: POKE 2001,143: POKE 2003,131: POKE 2004,136
58840  POKE 2005,129: POKE 2006,142: POKE 2007,135: POKE 2008,133
58845 JI(5) = 999: RETURN
58900  REM  POKE ?
58905  GOSUB 58000
58910  POKE 2012,191
58920  REM  POKE  # OF FILE =
58930  POKE 2000,163: POKE 2002,143: POKE 2003,134
58940  POKE 2005,134: POKE 2006,137: POKE 2007,140
58950  POKE 2008,197: POKE 2010,189
58990  RETURN
59000  REM
59010  GOSUB 58000
59020 JM = 19:JL = 28440
59030  GOSUB 58050
59035  IF TG(23) = 0 GOTO 59080
59040  FOR I = 1 TO  PEEK (25000)
59050  POKE 864, PEEK (25028 + I * 2)
59060  POKE 865, PEEK (25029 + I * 2)
59062  CALL 866
59064  FOR JK = 1 TO 5
59065  NEXT JK
59070  NEXT I
59080  RETURN
59100  REM
59110 JJ =  INT (TG(8) / 100)
59120  POKE 1622,JJ + 176
59130 JK =  INT (TG(8) / 10) - JJ * 10
59140  POKE 1623,JK + 176
59150 JK = TG(8) - JJ * 100 - JK * 10
59160  POKE 1624,JK + 176
59170  RETURN
59200  REM
59210 JM = 39:JL = 28420
59220  GOSUB 58050
59230  GOTO 51790
59300  REM
59310 JM = 39:JL = 28380
59320  GOSUB 58050
59330  GOTO 51790
59400  REM
59410  GOSUB 58000
59420 JM = 30:JL = 28340
59430  GOSUB 58050
59440  RETURN
59500  REM
59510 JM = 39:JL = 28310
59520  GOSUB 58050
59530  RETURN
59600  REM
59610 JK =  INT (TG(7) / 100)
59620  POKE 1874,JK + 176
59630 JJ =  INT (TG(7) / 10) - JK * 10
59640  POKE 1875,JJ + 176
59650 JJ = TG(7) - 100 * JK - 10 * JJ
```

```
59660  POKE 1876,JJ + 176
59670  RETURN
59700  REM
59710  GOSUB 58000
59720 JM = 29:JL = 28270
59730  GOSUB 58050
59740  RETURN
59800  REM
59810 JM = 39:JL = 28240
59820  GOSUB 58050
59830  RETURN
59900  REM
59910 JK =  INT (TG(9) / 100)
59920  POKE 1746,JK + 176
59930 JJ =  INT (TG(9) / 10) - JK * 10
59940  POKE 1747,JJ + 176
59950 JJ = TG(9) - 100 * JK - 10 * JJ
59960  POKE 1748,JJ + 176
59970  RETURN
60000  REM
60010 B$ =  STR$ (TG(21))
60020  PRINT D$;"BLOAD GRAPHIC ";B$
60040 JJ =  PEEK (28000)
60044 TG(20) = JJ
60046  IF TG(14) > TG(20) THEN TG(14) = TG(20)
60050 IM(0) = 27998 - JJ
60060  FOR QZ = 1 TO JJ
60070 IM(QZ) = IM(QZ - 1) - 3 * ( PEEK (27999 - QZ))
60080  NEXT QZ
60090  COLOR=  PEEK (27999)
60100  FOR JJ = 0 TO 39
60110  HLIN 0,39 AT JJ
60120  NEXT JJ
60130  RETURN
60200  REM
60210  FOR JJ = IM(QZ - 1) TO ((IM(QZ)) + 1) STEP  - 3
60220  POKE (256 * ( PEEK (JJ)) +  PEEK (JJ - 1)),( PEEK (JJ - 2))
60230  NEXT JJ
60240  RETURN
60300  REM
60310  FOR JJ = IM(QZ - 1) TO ((IM(QZ)) + 1) STEP  - 3
60320  POKE (256 * ( PEEK (JJ)) +  PEEK (JJ - 1)),( PEEK (27999))
60330  NEXT JJ
60340  RETURN
60400  REM  DISPLAY TOGGLED VALUES
60410  FOR I = 1  TO 5
60420  POKE (1760 + I),(I + 224 - TG(I) * 224)
60430  NEXT I
60460  POKE 1751,(245 - TG(6) * 224)
60470  POKE 1898,(24 + TG(18) * 224)
60480  POKE 1779,(6 + TG(19) * 224)
60490  POKE 1910,(238 - TG(23) * 224)
60499  RETURN
```

APPENDIX A6

EXPERIMENTAL DATA

A6-a	Phase 1 - Print out of test run.
A6-b	Phase 2 - Results of 64 test runs.
A6-c	Effects of error rate.
A6-d	Sounds used during phase 3 - experiment one.
A6-e	Sounds used during phase 3 - experiment two.
A6-f	Phase 3 - typical results sheet.
A6-g	Phase 3 - experiment one - summary of results.
A6-h	Phase 3 - experiment two - summary of results.

A6-a Phase 1 - PRINT OUT OF TEST RUN

CURRENT REJECT THRESHOLD=115
NAME ; BH
DATE ; 16A85
TIME ; 1625
COMMENTS ; -

ALPHABET	PILOT	MORSE
A	A	A
B	B	B CM
C	C	C
D	D	D
E	E	E IM
F	F	F
G	G	G CM
H	H	H
I	I	I EM
J KA	J	J
K	K	K
L	L	L
M NA	M	M
N	N	N
O	O	O
P	P	P
Q	Q	Q OM
R	R	R LM
S	S	S HM
T CA	T	T
U	U	U
V	V	V
W	W	W
X	X	X KM
Y	Y NA	Y
Z	Z	Z CM
23	25	17

	ALPHABET	PILOT	MORSE
REPEATS	2	25	15
SPURIOUS	0	0	0

A6-b PHASE 1 - RESULTS OF 64 TEST RUNS

ALPHABET VOCABULARY

REJECT THRESHOLD	EXPERIMENTAL DATA FROM FIVE TEST RUNS (UPPER = CORRECT / 26; LOWER = No. REPEATS)					AVG % (UPPER) / 0 (LOWER)	STD DEV 0	95 % CONFIDENCE LIMIT %	
10	24	24	20	18	17	79	3.2	61.2	97.2
						0			
20	22	20	22	21	22	82	0.8	77.4	87.2
						0			
30	15	21	17	21	17	69	2.6	55.3	84.6
						0			
40	23	21	21	20	22	82	1.1	76.0	88.5
						0			
50	18	22	15	18	20	71	2.6	57.2	85.8
						0			
60	21	20	22	21	19	79	1.1	72.9	85.4
						0			
70	22	18	19	19	17	73	1.8	62.8	83.3
						0			
80	17	23	19	19	16	72	2.6	57.6	86.9
						0			
90	22	19	20	20	20	77	1.0	71.6	83.6
						0			
100	21	21	17	17	21	74	2.1	62.6	86.6
						0			
105			21	20	20	78	0.5	75.7	80.6
				1		0.			
110	22	24	21	23	20	84	1.5	75.9	93.2
			2	2		0.			
115	24	21	21	18	23	82	2.3	69.7	94.9
	0	7	7	10	2	5.			
120	23	20	24	25	19	85	2.5	71.2	99.5
	33	52	34	43	91	50			

I.C.A.O VOCABULARY

REJECT THRESHOLD	EXPERIMENTAL DATA FROM FIVE TEST RUNS (UPPER = CORRECT / 26; LOWER = No. REPEATS)					AVG % (UPPER) / 0 (LOWER)	STD DEV 0	95 % CONFIDENCE LIMIT %	
10	25	23	22	21	23	87	1.	79.5	95.8
						0			
20	22	25	23	23	22	88	1.	81.7	95.1
						0			
30	21	23	25	23	22	87	1.	79.5	95.8
						0			
40	25	23	25	14	24	85	4.	59.8	110.
						0			
50	24	25	19	21	25	87	2.	73.0	102.
						0			
60	24	24	22	20	22	86	1.	76.9	95.3
						0			
70	25	25	22	21	24	89	1.	80.0	99.9
						0			
80	24	25	20	21	23	86	2.	75.5	98.2
						0			
90	26	26	19	25	24	92	2.	76.3	108.
						0			
100	26	26	25	24	22	94	1.	85.4	103.
						0			
105			20	21	24	83	2.	74.5	92.1
				2	3	1.			
110	26	25	25	24	25	96	0.	92.2	100.
			6	11	1	3.			
115	26	24	24	26	25	96	1	90.6	101.
		4	36	19	25	16			
120	26	26	25	25	24	96	0.	92.3	101.
	90	50	77	138	133	97			

MORSE VOCABULARY

REJECT THRESHOLD	EXPERIMENTAL DATA FROM FIVE TEST RUNS (UPPER = CORRECT / 26; LOWER = No. REPEATS)					AVG % (UPPER) / 0 (LOWER)	STD DEV 0	95 % CONFIDENCE LIMIT %	
10	18	13	17	12	16	58	2.	44.2	72.6
						0			
20	14	13	17	14	15	56	1.	47.8	64.4
						0			
30	17	16	13	12	13	54	2.	42.7	66.4
						0			
40	12	17	17	15	17	59	2.	48.0	71.9
						0			
50	13	15	13	9	16	50	2.	36.0	65.4
						0			
60	16	16	15	10	15	55	2.	41.6	69.1
						0			
70	18	21	14	17	9	60	4.	35.8	85.6
						0			
80	18	21	10	17	15	62	4.	39.9	84.6
						0			
90	15	22	11	15	12	57	4.	34.1	81.2
						0			
100	17	20	15	10	11	56	4.	33.3	78.9
						0			
105			14	11	18	55	3.	40.2	70.0
						0			
110	22	19	11	14	16	63	4.	39.6	86.4
						0			
115	22	17	12	17	17	65	3.	46.0	84.7
				14	15	5.			
120	22	18	16	14	18	67	2.	51.4	83.9
	44	32	90	124	127	83			

A6-c EFFECT OF ERROR RATE

1. NO ERRORS

We are justified in enforcing good morals for they belong to all mankind; but we are not justified in enforcing good manners, for good manners always mean our own manners. G.K. Chesterton.

2. 95% ACCURACY

We are justified i& enforcing good mo&als for they belon& to all mankind; b&t we are not justi&ied in enforcing g&od manners, for go&d manners always m&an our own &anners. G.K. Ches&erton.

3. 90% ACCURACY

We &re justified i& enforcin& good mo&als &or they belon& to all m&nkind; b&t we are not & justi&ied in e&forcing g&od mann&rs, for g&od ma&ners always m&an ou& own &anners. G.K. Ches&erton.

4. 85% ACCURACY

We &re &ustified i& e&forcin& go&d mo&als &or they belon& to all &&ankind; b&t we &re not & just&&ied in e&forcing g&od man&&ers, for go&d ma&ners &lways m&an ou& own &a&ners. G.K. Che&&erton.

5. 80% ACCURACY

We &re &&stified i& e&forcin& go&& mo&als &or they belo&& to all &&nkind; b&t we &re no& & just&&ied in e&forcing g&od ma&&&rs, for go&d ma&&ers &lways m&&n ou& own &a&&ers. G.&. Che&&erton.

A6-d SOUNDS USED DURING PHASE THREE - EXPERIMENT ONE

	FILE NUMBER	PRIME SOUND	TYPE OF ERROR			
			PLACE	MANNER	V/UV	F.C.D.
	0	DOG	GOG	ZOG	TOG	DO
	1	BUYS	GUYS	ZIGHS	PIES	BUY
	2	NIT	BIT	ZIT	NID	NI
	3	RIG	RID	ZIG	PIG	RI
	4	FIST	CIST	HIST	VIST	FIS
	5	BIGHT	BIKE	BUYS	PIGHT	BUY
	6	HIP	SIP	HISS	ZIP	HI
	7	VEST	ZEST	BEST	FEST	VES
	8	PARK	KARK	SARK	BARK	PAR
	9	DAISY	MAZY	VAISY	TAISY	DAIS
	10	JAMP	GAMP	DAMP	CHAMP	JAM
	11	FORT	SHORT	BORT	VORT	FOR
	12	EAT	EAK	EAF	EAD	EA
	13	THINK	FINK	TINK	VINK	THING
BLENDING	14	GREAT	RATE	BREAT	GATE	GREY
VOWELS	15	ARK	OOK	EEK	ORK	AR
FINAL CONSONANT DELETION	16	NEAT	-	-	-	KNEE
VOICED/UNVOICED	17	PAT	-	-	-	BAT
MANNER	18	VAT	-	-	-	BAT
PLACE	19	TAT	-	-	-	KAT

APPENDIX A6-d The 20 sets of sounds used during experiment one of phase three.

A6-e SOUNDS USED DURING PHASE THREE - EXPERIMENT TWO

TYPE OF ERROR	FILE NUMBER	SOUNDS COL. 1	SOUNDS COL. 2
PLACE	0	PARK	KARK
	1	BUYS	GUYS
	2	JAMP	GAMP
MANNER	3	BUYS	ZIGHS
	4	HIP	HISS
	5	PARK	SARK
VOICED / UNVOICED	6	GREAT	GATE
	7	PAT	BAT
	8	NIT	NID
	9	BIGHT	PIGHT
	10	PARK	BARK
FINAL CONSONANT DELETION	11	BUYS	BUY
	12	PARK	PAR
	13	BIGHT	BUY
	14	NIT	NI
	15	HIP	HI
	16	GREAT	GREY
	17	NEAT	KNEE
	18	BIKE	BUY
	19	THINK	THING

A6-f PHASE 3 - TYPICAL RESULTS SHEET

2.5.3 VOICE ASSESSMENT LOGSHEET.

YOUR NAME B H DATE 25 DEC 86

YOUR ROLE — ASSESSING STUDENT'S SOUND ☑

— RECORDING COMPUTER'S ASSESSMENT ☑

SUBJECT'S NAME PETER McK

REJECT THRESHOLD 70

NOTE: REJECT THRESHOLD MUST STAY THE SAME THROUGHOUT SESSION.

O = PROMPTED, X = SOUND MADE, — = UNACCEPTABLE, ? = QUESTIONABLE

P McK Library

DISC. LETTER	FILE No	SOUND				
		1	2	3	4	5
A	0	Ø Ø Ø				Ø Ø Ø
	1	Ø Ø Ø				
		X				O
						Ø Ø
	2	θ θ θ				
						θ θ Ø
	3	Ø Ø Ø				~~θ~~
		X				O
						θ
	4	θ θ θ				Ø Ø Ø
	5	θ Ø Ø				Ø Ø Ø
	6	Ø Ø Ø				Ø Ø Ø
	7	θ θ O				X
						Ø Ø Ø
	8	Ø Ø Ø				Ø
		X X				O O

A6-g **PHASE 3 - EXPERIMENT ONE - SUMMARY OF RESULTS**

Sheet 1 of 2.

TYPE OF SOUND	PRIME				PLACE				MANNER				V/UV				F.C.D.			
REFERENCE PATTERNS FROM	S	S	J	J	S	S	J	J	S	S	J	J	S	S	J	J	S	S	J	J
SPOKEN BY	S	J	J	S	S	J	J	S	S	J	J	S	S	J	J	S	S	J	J	S
FILE NUMBER 0	1	1			1	1			1						1					
FILE NUMBER 1	1		1	1	1				1	1	1	1	1	1		1	1	1		1
FILE NUMBER 2		1	1	1			1					1	1	1	1		1		1	
FILE NUMBER 3			1				1	1	1		1	1	1	1	1			1	1	1
FILE NUMBER 4					1		1		1	1					1				1	
FILE NUMBER 5	1	1		1	1				1		1				1		1	1	1	1
FILE NUMBER 6	1	1		1	1				1	1	1	1	1	1			1		1	1
FILE NUMBER 7						1			1		1	1			1				1	
FILE NUMBER 8	1	1		1	1	1				1	1	1		1	1		1	1		1
FILE NUMBER 9	1					1	1					1	1				1			
FILE NUMBER 10	1	1			1		1	1	1	1	1		1	1	1	1				
FILE NUMBER 11	1						1	1	1				1	1	1		1	1		1
FILE NUMBER 12	1		1			1	1			1		1					1	1	1	1
FILE NUMBER 13				1						1	1	1					1		1	1
FILE NUMBER 14	1	1	1	1		1	1					1	1	1		1		1	1	1
FILE NUMBER 15			1	1	1	1			1	1	1	1	1	1	1	1	1	1	1	
FILE NUMBER 16	3	3	3	2													3	2	3	3
FILE NUMBER 17	3	3	2	3													3	3	3	3
FILE NUMBER 18	2	3	2	3													3	1	1	0
FILE NUMBER 19	3	3	?	1													3	1	?	3
TOTAL - FILES 0 - 14	10	7	5	7	7	6	8	3	9	7	8	10	8	8	9	3	9	7	9	9

NOTES:

1. "1" INDICATES THE SOUND WAS RECOGNISED CORRECTLY ON THE FIRST ATTEMPT FOR FILES 0 - 15.
2. NUMBERS ASSOCIATED WITH FILES 16 - 19 REPRESENT NUMBER OF ENTRIES CORRECT OUT OF 3.

A6-g PHASE 3 - EXPERIMENT ONE - SUMMARY OF RESULTS

Sheet 2 of 2.

TYPE OF SOUND	PRIME				PLACE				MANNER				U/UV				F.C.D.			
REFERENCE PATTERS FROM	H	H	M	M	H	H	M	M	H	H	M	M	H	H	M	M	H	H	M	M
SPOKEN BY	H	M	M	H	H	M	M	H	H	M	M	H	H	M	M	H	H	M	M	H
FILE NUMBER 0			1		1	1	1		1	1			1	1		1	1	1	1	1
FILE NUMBER 1	1	1	1	1	1		1	1	1	1	1	1	1	1	1		1	1	1	1
FILE NUMBER 2	1			1					1		1	1	1		1		1		1	1
FILE NUMBER 3	1			1	1				1		1		1		1		1			
FILE NUMBER 4	1	1	1		1		1	1	1					1			1		1	1
FILE NUMBER 5		1	1		1		1	1	1	1			1	1	1	1	1		1	
FILE NUMBER 6	1			1	1		1		1			1				1	1	1		
FILE NUMBER 7		1		1			1		1			1		1	1					
FILE NUMBER 8	1	1	1		1		1	1	1				1			1	1			1
FILE NUMBER 9	1					1			1		1		1				1	1		1
FILE NUMBER 10	1	1				1	1	1			1	1	1		1		1	1	1	1
FILE NUMBER 11				1	1	1	1		1	1	1	1	1	1	1		1	1	1	1
FILE NUMBER 12				1				1	1				1			1				
FILE NUMBER 13	1	1	1						1							1	1	1	1	
FILE NUMBER 14				1			1	1	1				1	1	1	1	1	1		
FILE NUMBER 15	1		1	1	1		1		1	1		1	1	1	1	1	1	1	1	1
FILE NUMBER 16	3	3	3	2													2	0	2	3
FILE NUMBER 17	3	3	2	3													2	3	3	3
FILE NUMBER 18	2	2	3	3													3	2	2	3
FILE NUMBER 19	0	1	3	0													3	3	3	3
TOTAL - FILES 0 - 14	9	7	6	8	8	4	10	7	14	4	6	6	11	7	8	7	13	8	8	8

NOTES:

1. "1" INDICATES THE SOUND WAS RECOGNISED CORRECTLY ON THE FIRST ATTEMPT.
2. NUMBERS ASSOCIATED WITH FILES 16 - 19 REPRESENT NUMBER OF ENTRIES CORRECT OUT OF 3.

A6-h PHASE 3 - EXPERIMENT TWO - SUMMARY OF RESULTS

SOUND	COL. 1 OF APPENDIX A6-e											COL. 2 OF APPENDIX A6-e										
REFERENCE PATTERNS FROM	H	P	P	H	F	L	L	F	K	K	F	H	P	P	H	F	L	L	F	K	K	F
SPOKEN BY	H	H	P	P	F	F	L	L	F	K	K	H	H	P	P	F	F	L	L	F	K	K
FILE NUMBER 0	3	2	3	3	3	0	3	3	0	2	0	3	2	3	3	3	0	2	2	0	3	0
FILE NUMBER 1	3	0	3	3	3	0	1	0	0	2	0	1	2	2	0	0	0	2	0	0	2	2
FILE NUMBER 2	0	0	0	0	2	0	0	1	0	1	0	3	0	1	0	2	0	2	1	0	2	0
FILE NUMBER 3	3	0	3	0	3	0	0	0	0	3	0	3	0	0	0	1	0	3	0	0	3	0
FILE NUMBER 4	3	1	0	1	0	2	3	0	0	3	0	3	3	3	0	0	0	3	0	0	3	0
FILE NUMBER 5	3	1	2	0	3	0	0	0	0	2	0	3	0	3	1	1	0	2	0	0	2	0
FILE NUMBER 6	3	0	3	0	2	0	0	0	0	3	0	3	0	3	0	3	0	3	2	0	3	0
FILE NUMBER 7	3	0	0	0	0	0	1	0	0	2	3	3	0	3	0	0	0	3	0	0	3	0
FILE NUMBER 8	3	2	3	0	2	0	2	1	0	1	1	3	1	1	1	2	0	3	2	0	3	2
FILE NUMBER 9	3	0	3	0	3	0	3	0	0	1	0	3	0	0	0	2	0	1	0	0	3	0
FILE NUMBER 10	3	0	3	0	0	0	0	0	0	3	0	3	3	2	0	2	0	2	0	0	3	0
FILE NUMBER 11	3	0	3	0	0	0	3	0	0	1	3	3	0	3	0	2	0	3	2	0	3	0
FILE NUMBER 12	3	0	1	0	0	0	0	0	0	3	0	3	3	3	0	2	0	0	2	0	3	3
FILE NUMBER 13	2	0	3	0	0	0	3	0	0	3	0	3	0	3	0	3	0	3	0	0	3	0
FILE NUMBER 14	3	0	3	0	0	0	3	0	0	3	0	3	1	3	0	2	0	3	3	2	3	2
FILE NUMBER 15	3	0	3	0	1	1	3	2	2	3	1	3	1	3	3	3	1	3	3	2	3	2
FILE NUMBER 16	3	0	3	0	2	2	3	0	0	0	0	3	0	3	0	3	0	3	1	0	0	0
FILE NUMBER 17	3	1	3	0	3	0	3	2	0	2	0	3	0	3	1	3	2	3	3	3	3	3
FILE NUMBER 18	3	1	3	0	0	0	2	0	0	3	1	3	1	3	0	3	0	3	0	0	3	0
FILE NUMBER 19	3	0	3	3	0	0	2	0	0	2	0	3	0	3	2	3	0	3	1	0	2	0
TOTAL	56	8	48	10	27	5	35	9	2	43	9	58	17	48	11	40	3	50	22	7	53	14

NUMBERS REPRESENT NUMBER OF ENTRIES CORRECT OUT OF 3.

BIBLIOGRAPHY

(1) Adaptive Peripherals (1983) Adaptive Firmware Card Version 2e - Operators Manual. Adaptive Peripherals Washington.

(2) Ahrens, B.A., Wood, D.L., Hooper, J., Caudrey, D.J. and Seeger, B.R., (1985) Assisted Communication Needs of Physically Disabled Children. (Submitted for publication to the Journal of the International Society for Augmentative and Alternative Communication).

(3) Atal, B.S. and Hanauer, S.L., (1971) Speech Analysis and Synthesis by Linear Prediction of the Speech Wave, Journal of the Acoustical Society of America, Vol. 50, No. 2., pgs. 637-655.

(4) Baker, J.M., (1974) Time-Domain Analysis and Segmentation of Connected Speech, Speech Communications Seminar, Stockholm, pgs. 369-383.

(5) Birch, S., (July 1986) Verbal Istructions Can Pilot An Airliner, Aircraft, pg. 53.

(6) Bleck, E.E., (1978) Severe Orthopedic Disability in Childhood: Solutions Provided by Rehabilitation Engineering, Orthopedic Clinics of North America, Vol. 9, No. 2., pgs. 509-528.

(7) Boonzaier, D.A., Engelhardt, K.G. and Awad-Edwards, R. (1985) Interactive Evaluation of Current Voice Input Technology for the Physically Disabled Suggests New Avenues for Continued Research, RESNA 8th Annual Conference, Tennesee, pgs. 323-325.

(8) Bowe, F.G., Personal Computers and Special Needs, Sybex Computer Books.

(9) Brooks, P.L. and Frost, B.J., (1983) Evaluation of a Tactile Vocorder for Word Recognition, Journal of The Acoustical Society of America, Vol. 74(1), pgs. 34-39.

(10) Brooks, P.L., Frost, B.J., Mason, J.L. and Chung, K., (1985) Acquisition of a 250 Word Vocabulary Through a Tactile Vocorder, Journal of The Acoustical Society of America, Vol. 77(4), pgs. 1576-1579.

(11) Brown, C.C. and Cavalier, A.R., (1986) Increased Independence Through Computerised Voice Recognition for Persons Who are Severely Physically Involved and Profoundly Mentally Retarded, RESNA 9th Annual Conference, Minnesota, pgs. 101-103.

(12) Brown, M.K. and Rabiner, L.R., (1982) An Adaptive, Ordered, Graph Search Technique for Dynamic Time Warping for Isolated Word Recognition, IEEE Transactions on Accoustic, Speech and Signal Processing, Vol. ASSP-30, No. 4., pgs. 535-545.

(13) Caudrey, D.J., Seeger, B.R. and Walker, M.K. Technological Priorities for Enhancing Integration of Physically Disabled Children.

(14) Centre Industries, Myograph-Multipurpose Electromyogram Amplifier (Advertising Brochure).

(15) Clark, J.A. and Roemer, R.B., (1977) Voice Controlled Wheelchair, Archives of Physical Medicine and Rehabilitation, Vol. 58, pgs. 169-175.

(16) Creasey, G.H., Voice Operated Control for High Level Tetraplegia.

(17) Damper, R.I., (1982) Speech Technology - Implications for Biomedical Engineering, Journal of Medical Engineering and Technology, Vol. 6, No. 4., pgs. 135-149.

(18) Damper, R.I., (1984) Voice Input Aids for the Physically Disabled, International Journal of Man-Machine Studies, Vol. 21., pgs. 541-553.

(19) Damper, R.I. and Purswani, V.N., (1982) Voice Input Environmental Control for the Physically Disabled, World Congress on Medical Physics and Biomedical Engineering, Hamburg.

(20) Davis, K.H., Biddulph, R. and Balashek, S., (1952) Automatic Recognition of Spoken Digits, Journal of the Acoustical Society of America, Vol. 24(6), pgs. 637-642.

(21) Das, S.K., (1982) Some Experiments in Discrete Utterance Recognition, IEEE Transactions on Accoustics, Speech and Signal Processing, Vol. ASSP-30, No. 5, pgs. 766-770.

(21a) Doddington, G.R. and Schalk, T.B., (1981) Speech Recognition: Turning Theory to Practice, IEEE Spectrum, September 1981, pgs. 26-32.

(22) Frerich, R. and Brouwer, P. (1974) Spectral Analysis And Segmentation in Speech Recognition - New Developments and Experience, Speech Communication Seminar, Stockholm, pgs. 247-253.

(23) Fried-Oken, M., (1984) The Voice Recognition Device For Computer Control by Motorically - and Speech-Impaired Individuals, Proceedings of The 2nd International Conference on Rehabilitation Engineering, Canada, pgs. 454-455.

(24) Fried-Oken, M., (1985) Voice Recognition Device as a Computer Interface for Motor and Speech Impaired People, Archives of Physical Medicine and Rehabilitation, Vol. 66, pgs. 678-681.

(25) Fright, W.R., Bates, R.H.T., Brieseman, N.G., Garden, K.L., Kennedy, W.K., McCallum, B.C., Sewell, M.R., Tan, K.P. and Turner, S.G., (1985) Computer Translation of Impaired Speech, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, pgs. 72-77.

(26) FU, C., (1986) An Independent Vocational Workstation For a Quadriplegic, RESNA 9th Annual Conference, Minnesota, pgs. 182-184.

(27) Haton, J.P., (1974) Current Work in Segmentation and Analytical Recognition of Speech, Speech Communication Seminar, Stockholm, pgs. 197-200.

(28) Hendtlass, T., (1985) Speech Recognition and The Differently Abled, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, pgs. 32-35.

(29) IBM, (Winter 1986) IBM Scientists Demonstrate Personal Computer with Advanced Speech Recognition Capability, IBM Quarterly, pgs. 24-25.

(30) IBM (1986) Speech Recognition Research by IBM Scientists, IBM Bulletin June, pg. 7.

(31) Ignazzi, V. and Ramsden, V.S., (1984) Eye Operated Keyboard, Australasian Physical and Engineering Sciences in Medicine, Vol. 7, No. 2, pgs. 58-62.

(32) Iodesign Systems P/L., (1985) ESR-10 Speech Recognition Switch, (Advertising brochure).

(33) Jelinek, F., (1985) The Development of An Experimental Discrete Dictation Recogniser, Proceedings of The IEEE, Vol. 73, No. 11, pgs. 1616-1624.

(34) Kahn, D., Rabiner, L.R. and Rosenberg, A.E., (1984) On Duration and Smoothing Rules in a Demisyllable-Based Isolated-Word Recognition System, Journal of The Acoustical Society of America, Vol. 52(2), pgs. 590-598.

(35) Kaplan, G., Reddy, R. and Kato, Y., (1980) Words Into Action I,II,III, IEEE Spectrum, June, pgs. 22-29.

(36) Key Tronic Corporation, (1985) Speech Recognition in a Keyboard Peripheral, (Advertising Brochure).

(37) Klatt, D.H., (1977) Review of the ARPA Speech Understanding Project, Journal of The Acoustical Society of America, Vol. 62, No. 62, pgs. 1345-1366.

(38) Le Blanc, M., (1984) Rehabilitation Engineering Centre - With Research on Controls and Interfaces for Severely Disabled People, Final Report, 30.9.78 - 31.12.83.

(39) Levinson, S.E. and Liberman, M.Y., Speech Recognition by Computer, Scientific America, Vol. 244(4), pgs. 56-68.

(40) Levy, R., (1983) Interface Modalities of Technical Aides Used by People with Disability, American Journal of Occupational Therapy, Vol. 37, No. 11, pgs. 761-765.

(41) Lienard, J.S., Mlouka, M., Mariani, J.J. and Sapaly, J., (1974) Real Time Segmentation of Speech, Speech Communication Seminar, Stockholm, pgs. 183-187.

(42) Martin, T.B., (1976) Practical Applications of Voice Input to Machines, Proceedings of The IEEE, Vol. 64, No. 4, pgs. 487-501.

(43) Martony, J., Agelfors, E., Blomberg, M., Boberg, G., Elenius, K., Risberg, A., Spens, K.E. and Oster, A.M., (1974) Experiments with Electronic Lipreading Aids, Speech Communication Seminar, Stockholm, pgs. 93-96.

(44) Miller, J.D., Engebretson, A.M. and Defilippo, C.L., (1974) Preliminary Research with a Three Channel Vibrotactile Speech Reception Aid for the Deaf, Speech Communication Seminar, Stockholm, pgs. 97-157.

(45) Miller, G.E., Brown, T.E. and Randolph, W.R., (1985) Voice Controller for Wheelchairs, Medical and Biological Engineering and Computer, November 1985, pgs. 597-600.

(46) Microphonics, (1986) Pronounce (Advertising for a Speech Input System).

(47) Myers, C.S. and Levinson, S.E., (1982) Speaker Independent Connected Word Recognition Using a Syntax-Directed Dynamic Programming Procedure, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-30, No. 4, pgs. 561-565.

(48) Nirenberg, L.M., Hanley, J. and Stear, E.B., (1971) A New Approach to Prosthetic Control: EEG Motor Signal Tracking With An Adaptively Designed Phase - Locked Loop, IEEE Transactions on Biomedical Engineering, Vol. BME-18, No. 6, pgs. 389-398.

(49) Norman, D.A. (1983) Design Principles for Human Computer Interfaces, Proceedings of the CHI '83 Conference, Boston, pgs. 1-10.

(50) Olsen, H.F., Belar, H. and De Sobrino, R., (1962) Demonstration of a Speech Processing System Consisting of a Speech Analyser, Translator, Typer and Synthesizer, Journal of The Acoustical Society of America, Vol. 34, No. 10, pgs. 1535-1538.

(51) Olsson, C. and Walker, P.E., (1985) The Evolution of an Expressive Communication System, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, pgs. 49-52.

(52) Pickett, J.M., Gengel, R.W. and Quinn, R., (1974) Research with The Upton Eyeglass Speechreader, Speech Communication Seminar, Stockholm, pgs. 105-109.

(53) Poulton, A.S., (1983) Microcomputer Speech Synthesis and Recognition, Sigma Technical Press.

(54) Rahimi, M.A. and Eulenberg, J.B., (1977) Voice-to-Voice and EMG-to-Voice Systems for Speech Prosthesis, Proceedings of the National Electronics Conference, Vol. 31, pgs. 367-372.

(55) Raitzer, G.A., Vanderheiden, G.C. and Holt, C.S., (1976) Interfacing Computers for the Physically Handicapped - A Review of International Approaches, National Computer Conference, pgs. 209-216.

(56) Ramsden, V.S., (1985) The Development and Application of Communication Systems Using Optical Head Pointers, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, (Supplementary paper provided at Seminar).

(57) Reddy, D.R., (1967) Computer Recognition of Connected Speech, Journal of The Acoustical Society of America, Vol. 42, No. 2, pgs. 329-347.

(58) Richards, J.S., Wilson, T.L., Fine, P.R. and Rogers, J.T., (1982) A Voice-Operated Response Unit For Use In The Psychological Assessment of Motor Impaired Subjects, Journal of Medical Engineering and Technology, Vol. 6, No. 2, pgs. 65-67.

(59) Rivoira, S., (1974) Real-Time Automatic Detection of Syllable Nuclei in Continuous Speech, Speech Communication Seminar, Stockholm, pgs. 239-245.

(60) Rizer, B. and Hiner, L., (1985) Voice Activated Computing, Closing The Gap, Henderson, Minnesota.

(61) Rollins, A., Constantine, B. and Baker, S., Speech Recognition At Two Field Sites, Verbex, Bedford, U.S.A.

(62) Sakoe, H., (1982) Two Level D.P. Matching - A Dynamic Programing Based Pattern Matching Algorithm For Connected Word Recognition, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27(6), pgs. 588-595.

(63) Santech, P/L., "Some People Can Move Only Their Eyes", (Advertising brochure for Cedric Computer).

(64) Schmitt, G.D. and Tobias, J., (1986) Enhanced Communication For A Severely Disabled Dysarthric Individual Using Voice Recognition and Speech Synthesis, RESNA 9th Annual Conference, Minnesota, pgs. 304-306.

(65) Schwejda, P. and Vanderheiden, G., (1982) Adaptive-Firmware Card For The Apple II, BYTE, September, pgs. 280-317.

(66) Seeger, B.R., (1982) Aids For Handicapped Children, Crippled Childrens Association of South Australia Inc.

(67) Seeger, B.R. and Doble, W.D., (1980) The Development Of A Rehabilitation Engineering Service For Physically Handicapped Children, The Engineering Conference, Adelaide, pgs. 206-212.

(68) Seeger, B.R. and Hooper, J., (1985) Expressive Communication Technology - The Reality Beyond the Advertising, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, pgs. 83-93.

(69) Soede, M., (1986) Development of a Speech Recognition Command Set for Wheelchair Control, RESNA 9th Annual Conference, Minnesota, pgs. 104-106.

(70) Spine, T.M., Williges, B.H. and Maynard, J.F., (1984) An Economical Approach to Modelling Speech Recognition Accuracy, International Journal Man-Machine Studies, No. 21, 1984, pgs. 191-202.

(71) Sutter, E.E., (1983) An Oculo-Encephalographic Communication System, 6th Annual Conference on Rehabilitation Engineering, San Diego, 1983, pgs. 171-173.

(72) Sutter, E.E., (1984) The Visual Evoked Response as a Communication Channel, IEEE/NSF Symposium on Biosensors, 1984, pgs. 95-100.

(73) Traunmuller, H., (1974) A Visual Lipreading Aid, Speech Communication Seminar, Stockholm, Aug. 1-3 1974, pgs. 111-120.

(74) Vanderheiden, G.C., (1982) The Practical Use of Microcomputers in Rehabilitation, Bulletin of Prosthetic Research, Vol. 19, No. 1, Spring 1982, pgs. 1-5.

(75) Vanderheiden, G.C., (1982) Computers Can Play a Dual Role For Disabled Individuals, BYTE, Sept. 1982, pgs. 136-162.

(76) Voice Machine Communications Inc., Voice Input Module For Apple II, User's Manual, Voice Machine Communications, 1000 South Grande Avenue, Santa Ana, California, 92705.

(77) Voice Connection, (1985) The Voice Connection - With Only A Voice Command Everyday Things are Possible Again, CPL Services Pty. Ltd., 11-15 Wellington St., Windsor, Victoria, 3181.

(78) Waby, H., (1986) Bringing The Handicapped Into The Computer Age, Womens Day, Jan. 27 1986, pg. 28.

(79) Walpole, R.E., Myers R.H. (1972) Probability and Statistics for Engineers and Scientists. MacMillan Publishing Co., pgs. 190-192.

(80) Williams, C., (1985) Communication Aids: The Present Market in Australia, TADSEM '85: Australasian Seminar on Devices for Expressive Communication and Environmental Control, pgs. 11-21.

(81) Zue, V.W., (1985) The Use Of Speech Knowledge In Automatic Speech Recognition, Proceedings of The IEEE, Vol. 73, No. 11, Nov. 1985.

(82) Zwicker, E. and Terhardt, E., (1979) Automatic Speech Recognition Using Psychoacoustic Models, Journal of The Acoustical Society of America, Vol. 65, No. 2, Feb. 1979, pgs. 487-498.