THE DEVELOPMENT OF SYNTHETIC SPEECH AIDS FOR PATIENTS WITH ACQUIRED DISABILITY

by

© DAVID KEATING, C.PHYS., M.INST.P.

Senior Physicist,
Department of Clinical Physics and Bio-Engineering,
Glasgow

A thesis presented for the degree of Doctor of Philosophy
in the Faculty of Medicine, University of Glasgow

July 1988

TABLE OF CONTENTS

PAGE

APPENDICES

FIGURES		OPPOSITE PAGE

		OPPOSITE PAGE
A7.1	The Macleod Unit: Schematic Diagram	157
A7.2	The Macleod Unit: Switch Input Circuit	157
A7.3	The Macleod Unit: Power Up and Power Save Circuit	157
A7.4	The Macleod Unit: Power Supply Circuits	157
A7.5	The Macleod Unit: Memory Map	157
A7.6	The Macleod Unit: Address Decoding Circuit	157
A7.7	The Macleod Unit: Input/Output Memory Map	158
A7.8	The Macleod Unit: Input/Output Ports	158
A7.9	The Macleod Unit: Processor and Audio Circuit	159
A7.10	The Macleod Unit: Main Program Flow Chart	159
A7.11	The Macleod Unit: Code Entry and Verification Flow Chart	160
A7.12	The Macleod Unit: Speak and Display Words Flow Chart	160
A8.1	Text to Speech Mk I: Interface Memory Map	163
A8.2	Text to Speech Mk I: NVR Memory Map	163
A8.3	Votrax/Keyboard Buffer Interface	163
A8.4	Text to Speech Mk I: Power Supply Circuit	164
A8.5	Text to Speech Mk I: Main Program Flow Chart	164
A8.6	Text to Speech Mk I: Speak Routine Flow Chart	164
A8.7	Text to Speech Mk I: Store Data Flow Chart	164
A9.1	The Uvocom: Power Supply	165
A9.2	The Uvocom: Memory Map	165
A9.3	The Uvocom: Memory Address Decoding	165
A9.4	The Uvocom: Keyboard Circuit	166
A9.5	The Uvocom: Speech Processor and Audio Stage	167
A9.6	The Uvocom: Main Program Flow Chart	167
A9.7	Text to Phoneme Program Flow Chart	168

TABLES

OPPOSITE PAGE

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the encouragement and support of my supervisors - Professor J McKie, Director, West of Scotland Health Boards, Department of Clinical Physics and Bio-Engineering (DCPB) and Dr D J Wyper, Top Grade Physicist (DCPB).

I am greatly indebted to Dr A L Evans, Principal Physicist, DCPB for his suggestions and advice throughout and on the preparation of this thesis.

I would like to express my thanks to Mrs E Cunningham, Chief Speech Therapist, Southern General Hospital for her assistance and enthusiastic support.

Thanks for technical support has to be expressed to Mr R A Gowdie and Mr D C Smith, Medical Physics Technicians, DCPB, and also to Miss M Finlayson, Medical Physics Technician, DCPB, for her work on the illustrations.

I would like to thank Mr R McCrum, Mr R Wilson, Medical Physics Technicians (DCPB) and Miss K Gallagher, Speech Therapist, Southern General Hospital, Glasgow for their time and effort in recording the words and phrases for the speech devices.

Thanks to Miss Audrey Lawrence, Statistician, Southern General Hospital for her assistance with analysis of the intelligibility data and to Mr John McCall, Statistician, Glasgow University for his assistance with the analysis of the speech quality data.

Finally, I would like to thank my wife, Mrs Laura Keating, for typing the manuscript.

This work was financed by the Scottish Home and Health Department, Grant Number K/RED/4/C36.

DECLARATION

This thesis has been written entirely by myself. I had overall responsibility for the scientific component of the work. Technical assistance in the design and construction of the devices and statistical analysis of the speech quality trial data is detailed in the acknowledgements.

SUMMARY

Patients suffering from a variety of speech disorders can benefit from synthetic speech. This study concentrates on the dysarthric patients with acquired speech loss as these patients have intact intellect and are more likely to benefit from synthetic speech. The physical skills of these patients vary enormously and their needs and situations are different. The main part of this work is concerned with the design, development and evaluation of a range of speech aids to meet these varying needs and skills.

Three methods of speech synthesis are used and their performance has been investigated by using a Diagnostic Rhyme Test to measure the intelligibility of individual words. The results of this trial showed Adaptive Differential Pulse Code Modulation (ADPCM) to be more intelligible than Linear Predictive Coding (LPC), both these methods being more intelligible than constructive synthesis.

A further trial was conducted to measure the speech quality of phrases produced by the synthesisers. This showed listeners preferred listening to phrases constructed of LPC words than to phrases generated using Phoneme based synthesisers. Phrases with mixed LPC and constructed words were preferred to phrases of constructed words. The devices that were developed use different methods of synthesis and the choice of method was guided by these trials.

The Pocket Speech Aid is a rapid access limited vocabulary communication aid which uses ADPCM synthesis. Direct selection is the method used to give users access to eight phrases. The Pocket Speech Aid has been very successful in practice. When used as a telephone aid eight out of ten patients increased their communication ability and when used as a conversation prompter ten out of fourteen patients were able to steer the direction of real time conversations. This device has generated a great

deal of interest from other centres and the demand for the device which is currently being manufactured confirms that it has a role to play in assisting those with communication difficulties.

The Macleod Unit was named after a remarkable patient suffering from Motor Neurone Disease who realised his speech would soon be lost and had the foresight to select a vocabulary and record the words on a cassette recorder. His 625 word vocabulary was transferred to the speech aid which uses an encoding method of word selection. Clinical feedback showed the device to be of benefit for this highly motivated individual but was less successful for other patients in this group who found the cognitive effort to select codes too great.

An unlimited vocabulary device based on the commercially available VOTRAX which uses constructive synthesis was developed but this device was rejected because of the robotic sounding voice. A further unlimited vocabulary device prototype, the Uvocom, was designed to improve the speech quality and to investigate if there is a need for an unlimited vocabulary. The Uvocom uses a core vocabulary of 1000 LPC words and uses Phoneme back-up for words not stored in the core vocabulary.

Trials with the Uvocom have indicated that quality speech in an unlimited vocabulary device is likely to benefit a small number of patients who have the physical skills to operate such a device.

Finally, some indication is given of the directions in which future work could progress based on the proven success of the Pocket Speech Aid.

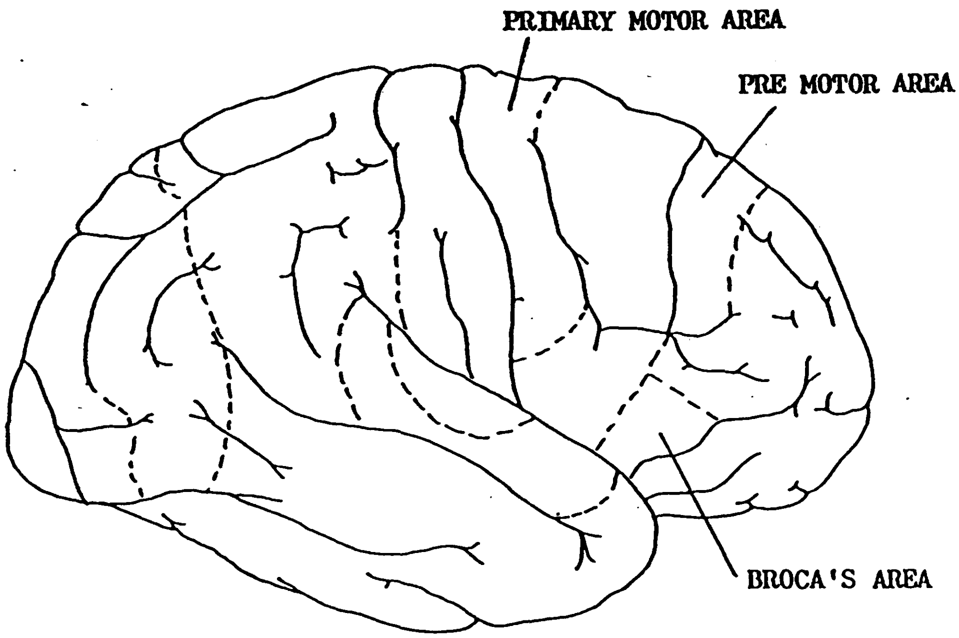

FIGURE 1.1 The Motor Areas of the Brain

CHAPTER 1 - INTRODUCTION

1.1 The Communication Process

Speech is the most powerful method of communication but it is not the only method. In everyday life, we use many other forms of communication such as facial expression, gesturing and writing. Although these other skills should always be encouraged when communication by speech is difficult, there are some situations where speech is essential.

The communication of a sound is the result of a complicated series of events. The first stage takes place in the brain and is the formulation of a concept at a linguistic or psychological level. The nervous system transmits this message to the speech organs and these behave in a conventional manner producing a particular sound. This second stage can be considered to be articulatory or physiological. The movements of the speech organs create disturbances in the air and these varying air pressures constitute the third stage in the chain, the physical or acoustic.

1.1.1 The Psychological System

Speech is a process contolled by functional areas of the cerebral cortex known as the motor areas of the brain (Fig 1.1). The primary motor area consists of regions that control specific muscles or group of muscles. The pre-motor area is concerned with learned motor activities of a complex and sequential nature. It generates impulses that cause a specific group of muscles to contract in a set specific sequence.

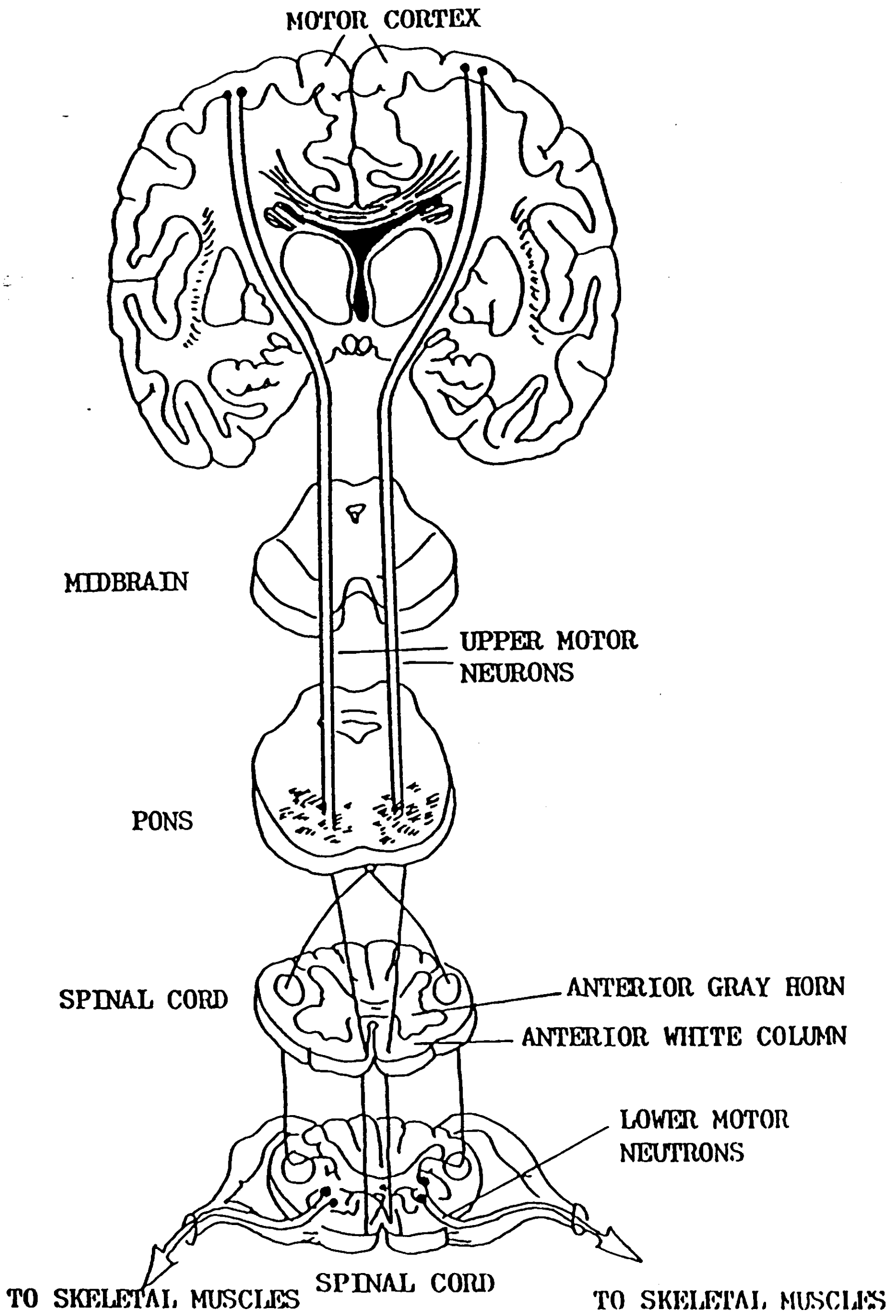

FIGURE 1.2 The Pyramidal Pathways

The translation of thoughts into speech involves Broca's area or the motor speech area. From this area, a sequence of signals is sent to the pre-motor regions that control the muscles of the larynx, the pharynx and the mouth. The impulses from the pre-motor area result in the specific coordinated contractions that are necessary for speech. Simultaneously, impulses are sent from Broca's area to the primary motor area and then onto the breathing muscles to regulate the proper flow of air past the vocal chords. The coordinated contractions of speech and breathing muscles enable the transfer of thoughts into speech. Voluntary movements are initiated by the brain and the signals travel through the spinal chord by two major routes, the pyramidal and extrapyramidal pathways (Fig 1.2). The pathways over which the impulses travel from the motor cortex to skeletal muscles have two components, upper motor neurones in the brain and lower motor neurones in the spinal chord.

The cerebral motor cortex plays an important role in skeletal muscle control but it is not the only area involved, the basal ganglia, the reticular formation, and cerebellum all contributing to muscle control. The motor cortex assumes the major role for controlling precise, discrete muscular movements. The basal ganglia largely integrate semi-voluntary movements like walking, swimming and laughing. The cerebellum, although not a control centre, assists the motor cortex and basal ganglia by making body movements smooth and coordinated.

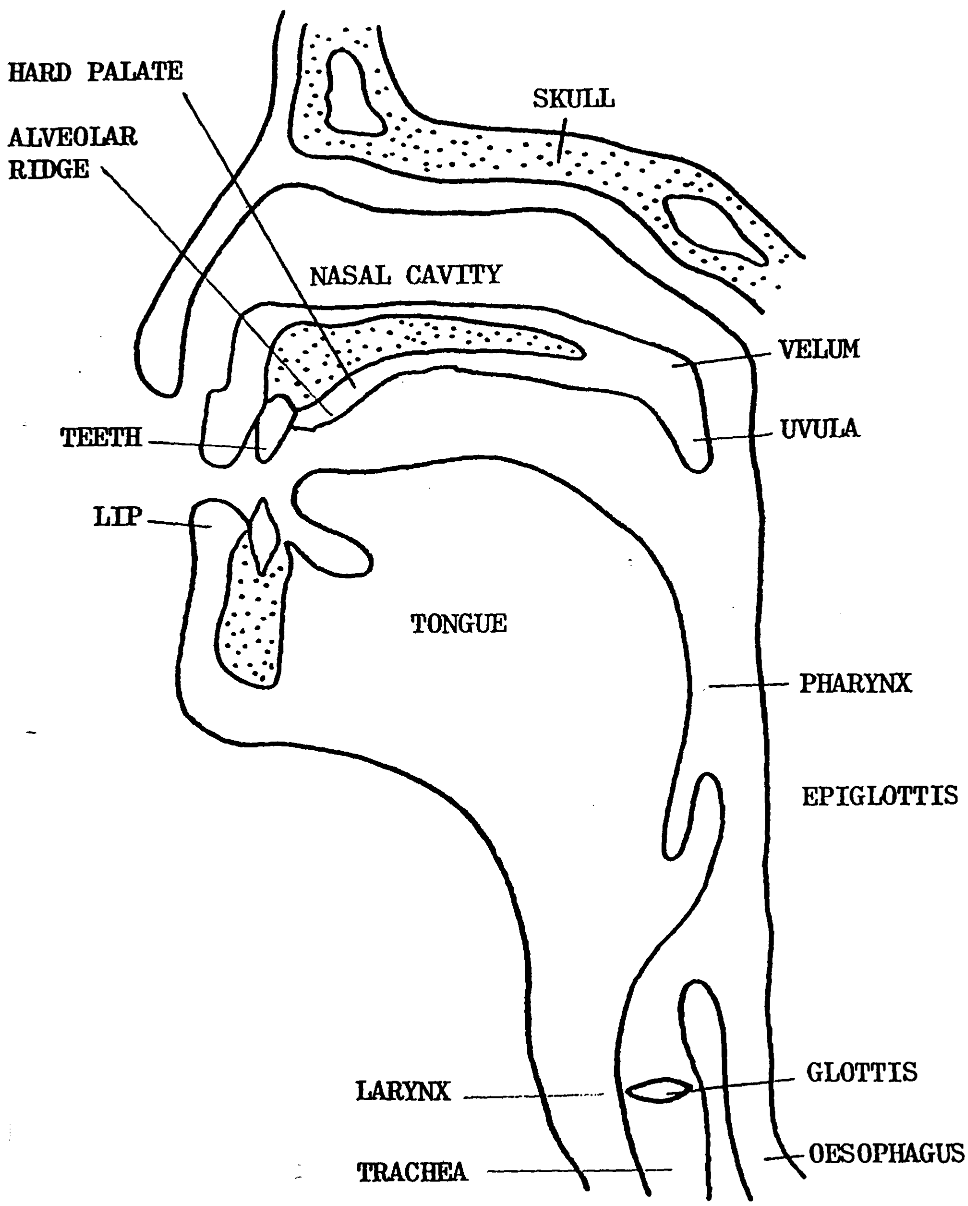

FIGURE 1.3 Human Speech Production Mechanism

1.1.2 The Physiological System

The mechanism of human speech production is shown in Fig 1.3. Four inter-related processes are involved.

a) **Respiration**

Energy for the speech process is provided by the respiratory system because speech is a pneumatic function depending on a moving column of air expired from the lungs. The air from the lungs moves upward through the bronchial tubes, trachea, larynx and pharynx and leaves the system through the oral or nasal cavity.

b) **Phonation**

Phonation is defined as the process by which sound is produced when air from the lungs sets into vibration the vocal folds which are enclosed in the larynx.

c) **Resonation**

Some components of the sound produced by the vocal folds can be amplified by altering the size and shape of the respiratory tract. The respiratory tract provides several resonators such as the larynx, the pharynx, the oral cavity and the nasal cavity.

d) **Articulation**

The air stream is shaped into fundamental units of speech known as phonemes by interruptions or obstructions caused by the articulators - the tongue, jaw, lips and teeth.

These four inter-related processes comprise the speech apparatus and they depend upon proper muscular contraction under neurone control. Each contraction must be made with exact force at an exact time and must be followed by other contractions in a precise succession. Each component must be contracted in exact relationship to the others. Speech disorders can result from a breakdown in any one of these four processes.

1.1.3 The Acoustic System

Speech sounds, like other sounds, are conveyed to our ears by means of compression and rarefaction of the air particles. These variations in pressure, initiated by the action of a vibrator, are propagated in all directions from the source, the air molecules vibrating at the same frequency as the original vibrator. In speech, these vibrations may be of a complex but regular nature producing 'tone', such as may be heard in vowel sounds; or they may be of an irregular kind, producing noise, such as in the consonant 's'; or there may be a combination of tone and noise, such as in 'z'. In the production of normal vowels the vibrator is the vocal chords and in the case of many consonants the source of air disturbance is provided by a constriction at a point in the vocal tract with or without accompanying vocal chord vibration.

The vocal chords vibrate in such a way as to produce, in addition to a basic vibration over their whole length known as the fundamental frequency, a number of harmonics which are multiples of the fundamental frequency. The number and strength of the component frequencies of this complex glottal tone will differ from one speaker to another. The quality of the tone is achieved by the shapes of the resonating chambers. These chambers are capable of an infinite number of shapes, each of which will have a characteristic vibrating resonance of its own. The harmonics of the glottal tone which coincide with the chamber's own resonance are considerably amplified. Thus certain bands of strongly reinforced harmonics are characteristic of a particular arrangement of the resonating chambers.

This complex range of frequencies of varying intensity which go to make up the quality of a sound is known as the acoustic spectrum. The bands of energy which are characteristic of a particular sound are known as the sound formants.

Chapter 2 will describe how speech synthesisers attempt to produce the sound formants by modelling the sound sources and the filtering characteristics of the vocal tract.

1.2 Communication Disorders

There are three major communication disorders which arise due to problems in the psychological or physiological systems.

The term aphasia means absence of language. Since few patients have complete absence of language the term dysphasia is used for impaired language function. Dysphasia is usually caused by damage to the frontal and temporal lobes of the brain and affects the four language modalities - auditory comprehension, reading comprehension, expression and writing - to differing degrees. Dysphasic patients have difficulty in the formulation and expression of intentions.

Anarthria is the term used if there is a complete loss of motor speech ability. Dysarthria comprises a group of speech disturbances due to partial loss of motor-sensory control.

Apraxia is the term used to denote a failure in voluntary programming of articulatory movements. Dyspraxia is a motor speech disorder which occurs due to left cerebral hemisphere injury. Dyspraxic patients have difficulty positioning their articulators but unlike dysarthrics suffer no impairment of muscle function. Dysarthria and dyspraxia are considered to be motor speech disorders and dysphasia a language disorder although many speech pathologists consider dyspraxia to have an element of language impairment.

Dysarthric patients often have intact comprehension and are more likely to benefit from a synthetic speech aid. Impairment of the neuromuscular events by neurological disease can affect all aspects of motor speech including respiration, phonation, resonance and prosody. The disease can involve different levels of the motor system and gives rise to five types of dysarthria (1).

Spastic dysarthria is a disorder of the upper motor neurone system which affects all voluntary movements including speech. The four major abnormalities of muscular function are spasticity, weakness, limitation of range and slowness of movement. The most prominent feature of disease of the upper motor neurone is the spasticity of the extensor muscles of the legs and in the flexor muscles of the arms.

Flaccid dysarthria is a disorder of the lower motor neurone system which affects all movement whether reflex, automatic or voluntary. The severity of impairment may vary from mild weakness to complete paralysis.

Extrapyramidal dysarthria or hypokinetic dysarthria are terms used for disorders of the extrapyramidal system. A marked limitation of the range of movement is the outstanding characteristic of hypokinesia. These patients suffer from rest tremor which produces visible joint movement.

Cerebellar or ataxic dysarthria is a disorder of the cerebellum. Voluntary movements are slow and the force, range, timing and direction of movements are inaccurate. Patients with severe cerebellar dysarthria may also suffer from extremity tremor.

If there is a disorder of multiple motor systems then the dysarthria is said to be mixed. The characteristics associated with each individual dysarthria are combined in mixed dysarthria.

It is not possible to state which of the dysarthric sub-types are more likely to benefit from a synthetic speech aid. The choice of device is heavily influenced by the stage of the disease and by the patient's physical disabilities. Some features such as tremor or range of movement impose restrictions on the type of switch input which can be used to drive a speech aid.

1.3 Conventional Speech Therapy

The development or redevelopment of functional speech may be possible for some patients suffering from motor speech disorders. These patients can benefit from a programme of speech rehabilitation. Therapy is aimed at teaching the patient to capitalise on the remaining motor potential.

1.3.1 Therapy for Dysphasia

Dysphasic therapy is aimed at treating the four language modalities - reading comprehension, auditory comprehension, expression and writing (2).

Therapy of auditory and reading comprehension would begin at an intervention level determined by detailed assessment. The earliest starting point could be with gesturing or at single word level, the material presented being gradually increased in complexity to increase comprehension ability. The level of intervention for expressive language treatment would also be determined by detailed assessment, the earliest starting point being at an automatic level such as counting days or months and progressing to completion of phrases and then answering questions and picture description. Writing skills would be improved in a similar manner.

1.3.2 Therapy for Dysarthria

A therapy programme for a dysarthric patient is aimed at strengthening movements of the tongue, lips and soft palate to increase speech intelligibility (3). The therapist can heighten awareness of position and feeling in weak areas through the use of icing sugar brushed round the mouth, lips and soft palate. Lip closure and tongue movements are encouraged by repetitive drilling on sounds, words and then longer utterances.

Because the muscles are weak, exercises in breath control and rate control are used in order to slow speech down and to give the muscles time to make the movements for the speech sounds.

1.3.3 Therapy for Dyspraxia

As dyspraxia normally occurs along with an element of dysphasia some of the techniques used in dysphasic therapy are applicable. These may be combined with sensory training. The therapist tries to achieve a consistent realisation of a sound through the use of sound drilling using sensory feedback. The goal here is to help the patient to regain some of the accurate voluntary control in programming the position of his articulators to produce the required phoneme sequences (4).

1.3.4 The Need for Communication Aids

Conventional speech therapy may not be successful in developing functional speech for patients with severe communication disorders. For progressive disorders speech therapy is useful in the early stages but as the disease progresses the therapist is involved in assessment for other forms of communication.

Writing, word boards, charts and symbol systems are used to enable these patients to communicate. In recent years electronic devices have become available and these have increased the range of assistive devices for use by the speech impaired. The use of synthetic speech in these assistive devices is the subject of this investigation.

The ideal speech aid would enable a speech impaired individual to communicate at a normal rate in all situations. In practice, this ideal speech aid is far from attainable but a practicable communication aid can be regarded as an assistive device which can maximise the communication potential of a speech impaired individual. A practicable speech aid can match some of the features of natural speech but it is not possible to have natural speech at a normal communication rate with unlimited vocabulary in any one device. A practicable speech aid can satisfy some of these features at the expense of others.

1.4 Patient Groups who Might Benefit from Using a Synthetic Speech Aid

There are no accurate figures for incidence and prevalence of patients in the United Kingdom with speech and language disorders. A recent review of the literature by Enderby and Philip (5) has estimated the number of people in the United Kingdom with communication disorders to be around 800,000. The following section on the patient groups likely to benefit from synthetic speech uses the incidence and prevalence figures quoted by Enderby and Philip.

1.4.1 Progressive Degenerative Diseases

Parkinson's Disease

The speech disorder associated with Parkinson's disease is extrapyramidal hypokinetic dysarthria. When speech is affected by Parkinsonism there is a reduced mobility, a restricted range of movement and a higher than normal rate of repetitive movement. The most prominent speech deviations associated with this disease are reduced phonation, volume and intonation and an increased rate of speech.

An estimated 36% of Parkinson patients have no speech difficulty. There is a prevalence rate of 150 per 100,000 population while the incidence rate is 20 per 100,000 population per year. There are an estimated 100,000 Parkinson's patients in the United Kingdom (6).

Multiple Sclerosis

This is an inflammatory condition involving mainly the white matter of the central nervous system. Some 20 - 40% of patients with multiple sclerosis experience dysarthria (7,8). The dysarthria is of mixed type, some speech deviations resemble that of ataxic dysarthria and other speech deviations that of spastic dysarthria. The most prominent speech deviations are impaired control of loudness and harshness.

The prevalence rate for multiple sclerosis is 60 per 100,000 and the incidence of the disorder is 3 per 100,000 population per year in the United Kingdom (9).

Fredreich's Ataxia

This is a disorder resulting from damage to the cerebellum. The speech of patients suffering from this disease is characterised by cerebellar dysarthria and the salient features are both inaccuracy and slowness of movement. In nearly all advanced cases of the disease the muscles of articulation are involved resulting in slurred speech.

The prevalence of this disease is 2 per 100,000 population (10) and there are an estimated 1,000 cases in the United Kingdom (8).

Myasthenia Gravis

Myasthenia Gravis is a disease of the lower motor neurone system. The speech disorder of patients suffering from this disease is that of flaccid dysarthria. A distinctive feature of myasthenia gravis is the progression and increase in severity of speech problems with prolongation of speech activity. Patients suffer from fatigue of the speech muscles.

The incidence rate of myasthenia gravis is 3 per 100,000 population per annum and there may be 2,800 sufferers in the United Kingdom. This estimate is based on prevalence rate studies in Finland (11).

Huntington's Chorea

Huntington's Chorea is a disease of the extrapyramidal system and patients with this disease exhibit hyperkinetic dysarthria. Voluntary muscle movements are invariably slow and there are also quick involuntary movements associated with the disease.

The incidence rate is 0.4 per 100,000 population per annum and the prevalence rate is 5 per 100,000 (12). A high proportion of sufferers are dysarthric.

Motor Neurone Disease

Motor neurone disease is characterised by a progressive degeneration of the neurons in both upper and lower motor neurone systems. The speech disorder associated with motor neurone disease patients is mixed dysarthria. As the patients have both bulbar and pseudobulbar palsy both spastic and flaccid dysarthria characteristics are present.

The incidence rate for this disease is estimated at 1.6 per 100,000 per year and a prevalence of 6 per 100,000 (13). 25% of patients show dysarthria and dysphagia as the first symptom (14) and almost all will eventually have bulbar involvement and associated flaccid dysarthria.

1.4.2 Non-Degenerative Diseases

Cerebrovascular Accident (CVA)

An estimated 2 persons per 1,000 population suffer a cerebrovascular accident each year. The survival rate is between 63 - 77% (15,16). Approximately one third of the surviving group will have severe communication difficulty. Dysphasia is the most common disorder but dysarthria can also be present. Many stroke patients recover their speech within six months.

Head Injury and Neurosurgery

In 1972 in England and Wales there were 142,000 admissions to hospitals of patients suffering from the effects of head injury (17). 15% of these patients experienced speech disorders. Some studies suggest that as many as one third of survivors have speech disorders many years after head injury.

1.4.3 The Potential Number of Users

In the United Kingdom there are an estimated 140,000 patients with progressive degenerative diseases and an estimated 700,000 patients with non-degenerative diseases. Enderby and Philip have estimated that 20,000 of those with progressive degenerative diseases and 70,00 of those with non-degenerative diseases have severe communication disorders (little or no spontaneous speech possible). Many of these patients could benefit from the use of a synthetic speech aid.

1.5 The Problems in Prescribing a Speech Aid

Although the technology now exists to enable low cost portable speech aids to be developed, there remains the considerable problem of identifying the most suitable aid for each client. The range of physical and cognitive skills is enormous. Other factors such as the patient's needs and the situations in which the device is to be used further complicate the problem of prescription. The development of communication aids has enabled some individuals previously lacking the power of self expression to produce items of prose which reveal extraordinary intellectual ability (18).

1.5.1 Physical and Cognitive Skills

The range of physical skills for patients suffering from acquired speech loss is considerable. Some patients are still mobile whereas others might only be able to operate a single switch. Cognitive abilities also cover a wide range but this investigation concentrates on patients with acquired speech loss who often have intact intellect.

1.5.2 Needs and Situations

If the patient has many communication partners in a variety of situations then the communication aid may have to be of a general nature or more than one device may be required. However, in some cases the device may only be required for use in specific situations and a device suited to these situations could be of more benefit. Each case must be treated on its own merits when considering these factors.

1.6 The Development and Evaluation of a Range of Synthetic Speech Aids

This work describes the development and evaluation of a range of speech aids. The range of devices is designed to meet a variety of patient needs and situations. The devices are also designed to meet a wide range of physical disabilities. This section investigates some of the technical features in the development of the devices.

1.6.1 The Input Mechanism

A speech aid should try to maximise the communication rate of a disabled individual. A major limiting factor of the communication rate is the selection mode for a device. There are three selection techniques used for interfacing to a communication aid.

a) **Direct Selection**

This is cognitively the most simple method for interfacing but it is physically the most difficult of the selection methods. The total number of vocabulary items are displayed to the patient and selection of a letter, word or phrase is achieved by pointing directly to the desired item as in the case of conventional word or alphabet boards, or by pressing a single switch associated with a particular item. The number of switches in a direct selection device is equal to the number of vocabulary items present. Direct selection is the fastest selection method but requires good physical control.

b) **Scanning**

A scanning technique is used if a patient has a limited amount of physical control. Scanning methods only require the operation of one or two switches to control a communication aid and so the number of switch inputs is much smaller than the vocabulary size. The communication rate is much slower than for direct selection because the patient must wait for the desired vocabulary item to become available for selection.

c) **Encoding**

Encoding can be used to increase the vocabulary size over the limits imposed by the number of switches in direct selection; this is achieved, however, at the expense of speed. Each vocabulary item is assigned a code which can be entered by pressing a small number of switches. Encoding requires both good cognitive skills and reasonable physical skills.

1.6.2 Range of Vocabulary

The vocabulary size is related to the method used for synthesising speech. Chapter 2 describes the different methods. In general, aids range from devices with a small fixed vocabulary of natural-sounding speech to those with unlimited vocabulary but with mechanical-sounding speech.

1.6.3 Ease of Operation

If the cognitive and physical effort required to operate a communication aid outweigh a user's motivation then eventually the device will be rejected. The physical effort can be minimised by suitable choice of input interface for a particular patient's disability. Scanning devices can be operated by any type of single switch e.g. push-button, suck-puff, brow-wrinkle or eye movement switch. Encoding devices can use different sizes of keys and key guards. The direct selection devices can be physically small or large. If it requires less effort to communicate by some other means then this will be the method a user will choose to communicate.

1.6.4 **Portability**

Each user of an augmentative communication system will have a different set of needs and a different set of situations requiring the use of a communication aid. For some patients it may not be neccessary to have a portable device as the communication system may only be required to be used in one specific situation such as in an intensive care unit or in a hospital ward. However, for able bodied dysarthric patients who are likely to want to communicate in a variety of situations, portability is an important feature for a device. It enables them to enjoy a greater degree of freedom if they are not restricted to the use of their device in one specific situation.

1.6.5 **Speech Quality**

Intelligibility and the degree of naturalness are two features which are important to the overall speech quality of a device. The intelligibility and the degree of naturalness of the speech produced by the different methods of synthesis are examined in detail in chapter 3.

1.6.6 **Cost**

Cost is an important factor which limits the number of devices reaching the patients. Unlimited vocabulary synthesisers with high quality speech are available but only at a cost of over a thousand pounds.

1.7 Summary

This chapter has shown that breakdown in the physiological systems of human speech production can give rise to several different types of communication disorder. Although dysarthric, dysphasic and dyspraxic patients can all benefit from synthetic speech this study concentrates on the dysarthric group. The particular group of patients studied are dysarthric patients with acquired speech loss. These patients have intact intellect and are therefore more likely to benefit from the use of synthetic speech. The physical skills of these patients vary enormously and their needs and situations are different. This work describes the design, the development and the evaluation of a range of synthetic speech aids for those patients with acquired speech loss.

CHAPTER 2 - SPEECH TECHNOLOGY

2.1 Introduction

Production of high quality synthetic speech has been possible for over twenty years, but only on large computer systems costing several thousand pounds. Recent advances in microelectronic technology has enabled the addition of synthetic speech to a commercial product for little extra cost (19,20,21).

Inventors have been producing speaking machines since the 18th century. A Hungarian engineer, Von Kempelen, built one of the earliest mechanical speaking machines which was difficult to operate and required the setting of numerous stops and levers, but it could produce almost any English sound. Alexander Graham Bell also constructed a mechanical talker in 1922; this device could even produce simple phrases. In 1940 the Bell laboratories developed the spectrograph which for the first time allowed the analysis of speech waves and illustrated their variation with time. A few years later vocoders were developed. These were devices which analysed a speech signal, measured the fundamental pitch and the amplitude in different frequency bands and determined whether or not the signal was periodic. These parameters were then used to reconstruct the original signal. This was the real starting point for modern speech synthesisers.

In the late sixties high quality speech could be produced using formant synthesisers developed by Fant and others at the Speech Transmission Laboratory in Sweden. The OVE 3 synthesiser was controlled by a CDC 1700 digital computer, the synthesiser parameters, formant frequencies, amplitudes and bandwidths being controlled by the computer through digital to analog converters. Formant synthesis is essentially an analog technique as it requires multiple filters to represent the filtering action of the vocal tract.

In 1971 a new technique known as Linear Predictive Coding (LPC) was developed which was ideally suited to digital techniques. At this time high density memory devices were being developed allowing more and more speech data to be stored in read only memory (ROM) devices. By 1977 speech began to appear in commercial products such as the Talking Calculator by Telesensory Systems Inc and Speak 'n' Spell by Texas Instruments, both these devices used the new method of Linear Predictive Coding. In 1978 it would have been possible to store only 8 seconds of LPC speech in the largest available programmable device, the 2708 (1Kbyte PROM). In 1987, 17 minutes of speech can be stored in the largest readily available programmable device the 27C1024 (1Mbyte PROM). Digital signal processors in single chip form have now been developed and dedicated processors capable of performing the complex LPC algorithm for reconstruction of the speech signal have become available allowing more complex processing to be done in real time.

2.2 The Methods of Synthesising Speech

The three main methods used for synthesising speech are waveform coding, synthesis by analysis and constructive synthesis. These modern methods are digital and differ in memory requirements, speech quality and hardware complexity. When converted to an electrical signal by a microphone, human speech is a continuously variable analogue signal which must be digitised to be stored in electronic memory devices. To obtain a digitised representation of a continuous analogue signal the signal is first sampled to convert it to a series of discrete analogue samples and these samples are converted to digital form. Shannon's sampling theorem (22) states that a signal can be completely defined if the sampling rate is greater than twice the maximum frequency present. To digitise human speech, the sampling rate must be at a rate greater than twice the maximum frequency present which for telephone quality speech is about 4KHz. A sampling rate of 8KHz would therefore be required and if this were to be digitised to 8 bit accuracy then it would require 64,000 bits of information to store 1 second of speech. This is the minimum bit rate necessary for digitised speech but in practice 12 bit accuracy is used giving a bit rate of 96,000 bits/sec. This is the technique known as Pulse Code Modulation (PCM) used for converting speech to digital form in telecommunications (23). The bit rate is high but the hardware complexity and the costs are low requiring only standard digital to analogue converters to reconstruct the original speech signal.

Modern speech synthesis methods try to reduce the memory requirements for the digitised data (24). The different forms of these three methods will be described together with an outline of the particular techniques used in this project.

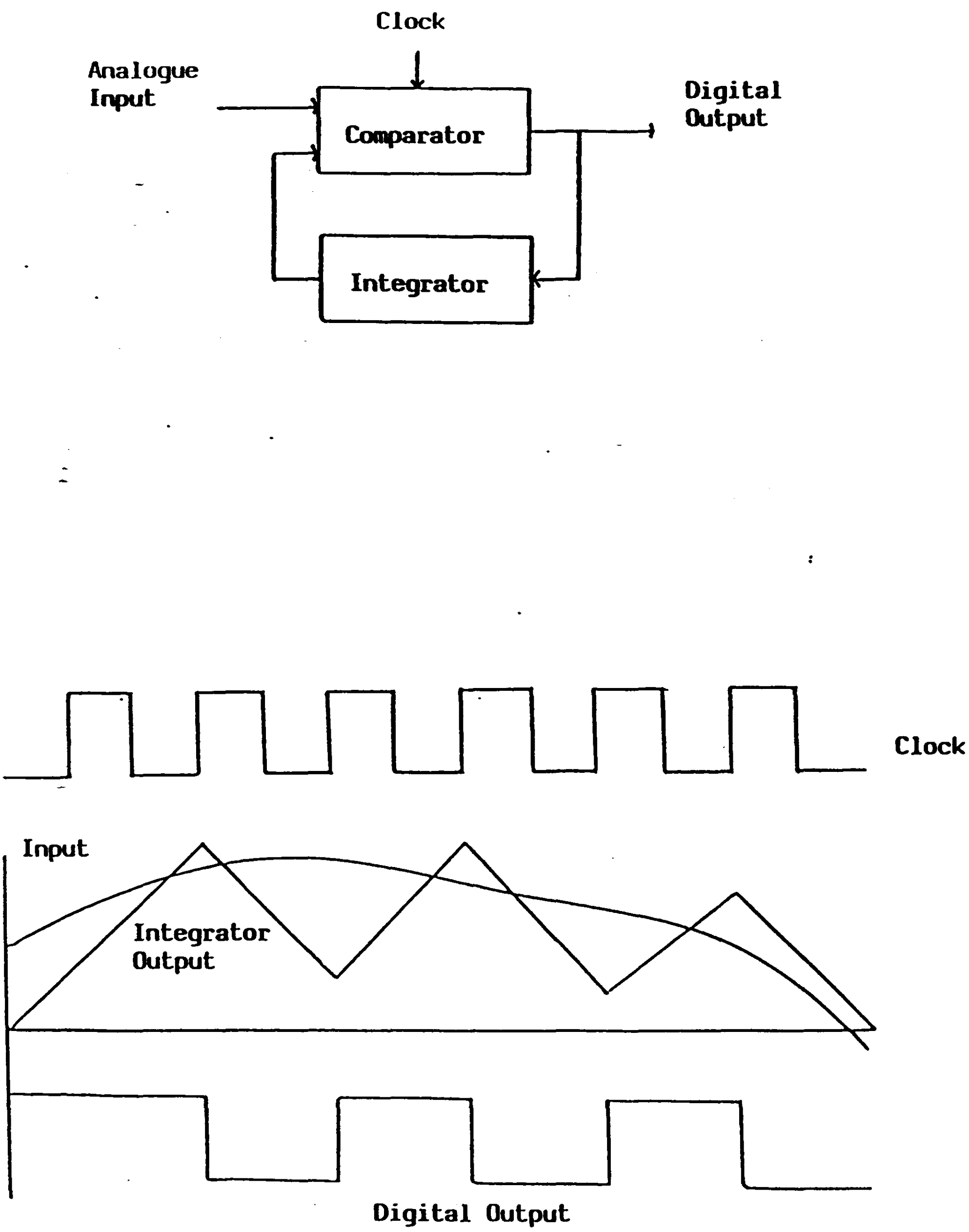

FIGURE 2.1 The Delta Modulator

2.3 Adaptive Differential Pulse Code Modulation (ADPCM)

ADPCM is an example of waveform coding. The technique of ADPCM synthesis was derived from the techniques of delta modulation and pulse code modulation, (25, 26).

A delta modulator stores only the difference between consecutive samples rather than the absolute values. In its simplest form the delta modulator uses only one bit per sample and merely indicates whether the incoming signal is increasing or decreasing.

Using delta modulation to digitise a speech signal requires only very simple hardware. Figure 2.1 shows a block diagram of a delta modulator. The integrator produces a ramp voltage which can be either increasing or decreasing. The incoming signal is compared with the integrator output and if the signal is more positive than the integrator then a 1 bit is stored and the integrator is set to ramp in the positive direction and if the signal is more negative than the integrator then a 0 bit is stored and the integrator is set to ramp in the negative direction. To reconstruct the signal the digital data is simply passed through an integrator to give the speech output. The use of only 1 bit per sample requires a very high sampling rate if the digital to analogue converter is to follow the incoming signal successfully. If the signal changes to a higher frequency and a higher amplitude then the integrator slope would be unable to follow the incoming signal.

Refinements on the basic delta modulator make use of the fact that the intrinsic feedback loop gives the circuit a form

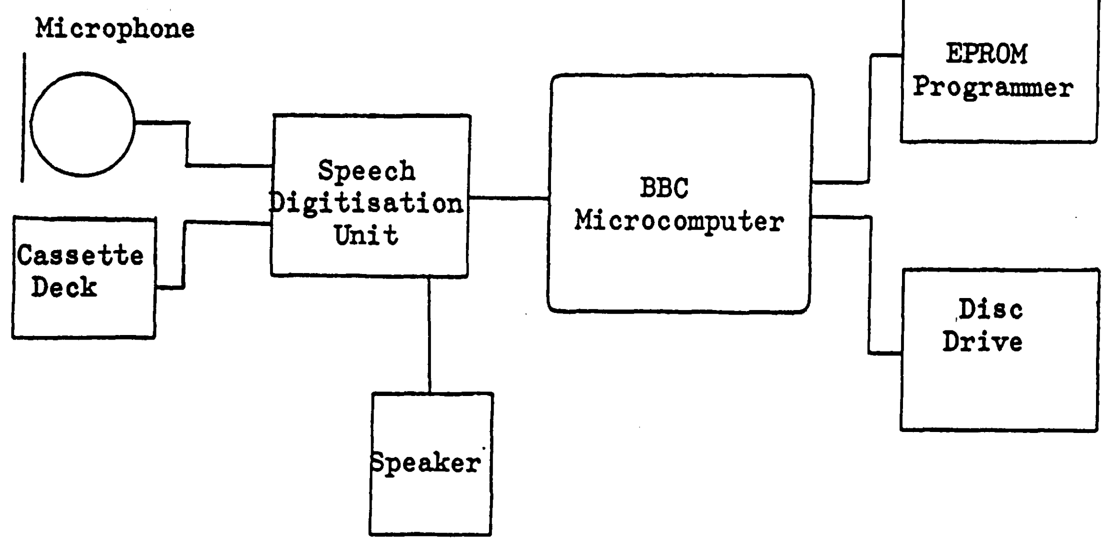

FIGURE 2.2 The ADPCM Recording System: Schematic

of memory. If several 1 bits in succession have been stored or transmitted the it is likely that the integrator is not ramping fast enough. The advantage of delta modulation can be combined with pulse code modulation if instead of 1 bit quantisation a PCM coder is used in the feedback loop. This technique is called Differential Pulse Code Modulation (DPCM).

The uncertainty present in the digital sampling process of PCM is known as quantisation noise. During loud speech the noise might not be noticeable but during low level speech the quantisation noise can become significant.

Adaptive techniques allow the noise to vary by making the quantisation step change as the signal amplitude changes. Adaptive pulse code modulation uses the basic technique of DPCM with an adaptive quantisation value which changes with variations in the amplitude of the input waveform (27).

Recording of ADPCM Data

Figure 2.2 is a block diagram of a system which has been developed to record ADPCM data. Analogue signals can be sent to the input pre-amplifier from a microphone input or from a cassette tape. The interface unit (28) consists of an analogue to digital converter to convert signals to digital format for the OKI speech synthesiser chip (MSM5208). The synthesiser converts the input data to ADPCM format which is then sent to a BBC B microcomputer via the 1 MHz bus. The software running on the BBC microcomputer is responsible for sampling the data arriving via the 1 MHz bus and for storing the data in memory. The program allows for data to be replayed through the same interface unit because the 5208 synthesiser is an

FIGURE 2.3 Using the ADPCM Recording System

analysis/playback device. Data can be stored on disc or transferred to the EPROM programmer unit. The recording system is shown in Figure 2.3.

ADPCM synthesis is a method of waveform coding which requires only inexpensive recording hardware. The data rate can be externally set from 12 Kbits/sec to 32Kbits/sec. A 8KHz sampling rate and 4 bit coding were selected giving a data rate of 32K bits/sec in order that the speech quality is very high and human characteristics of the speakers voice are preserved.

2.4 Linear Predictive Coding (LPC)

2.4.1 Theory

LPC is a frequency domain method of synthesis by analysis. The human speech production mechanism has already been described in Section 1.2 and LPC exploits the fact that the mechanical inertia of the articulators implies a finite time for their displacement and therefore the vocal tract can be considered stationary over a small time interval. A human voice is digitised and then analysed by a computer to extract the energy, pitch and a set of vocal tract variables necessary to reconstruct the original signal. These parameters are calculated for each frame interval which is usually 20 msec and then stored in the computer.

For any input signal, a sample can be predicted from the sum of a large number of linear and higher order terms of previous samples. Modern predictive synthesisers use only linear

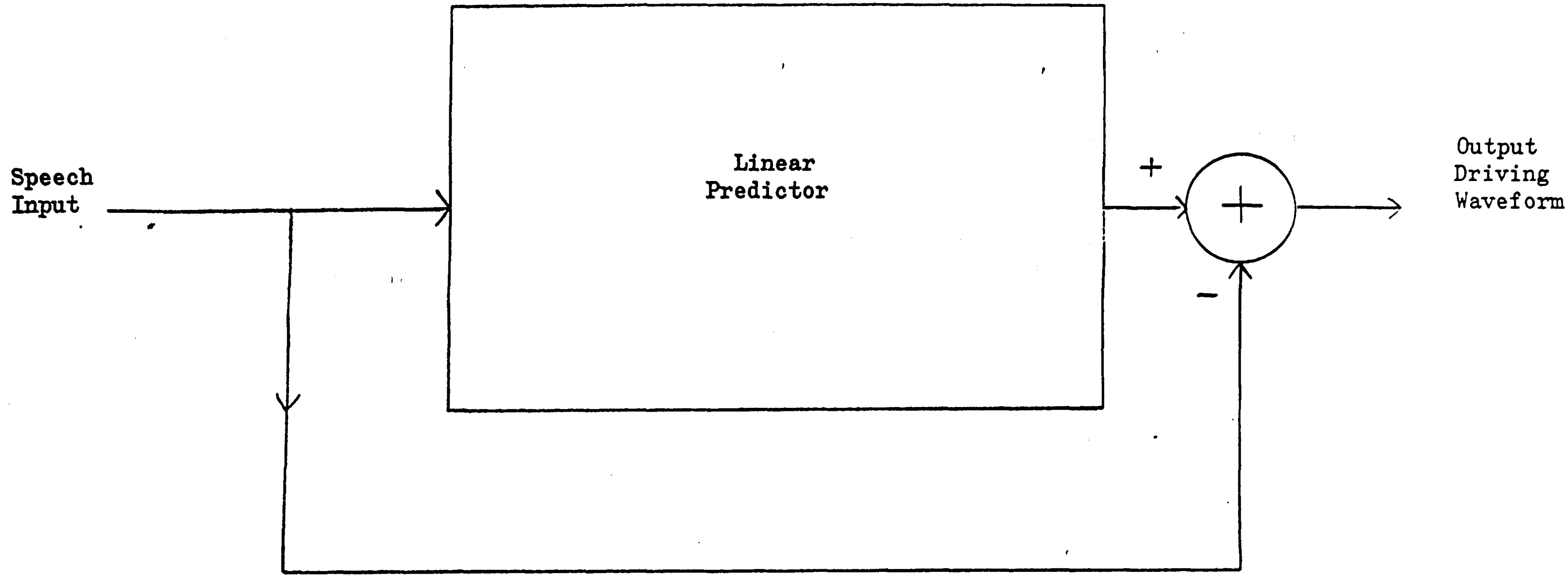

FIGURE 2.4 **Extraction of the Glottal Waveform by Linear Predictive Coding**

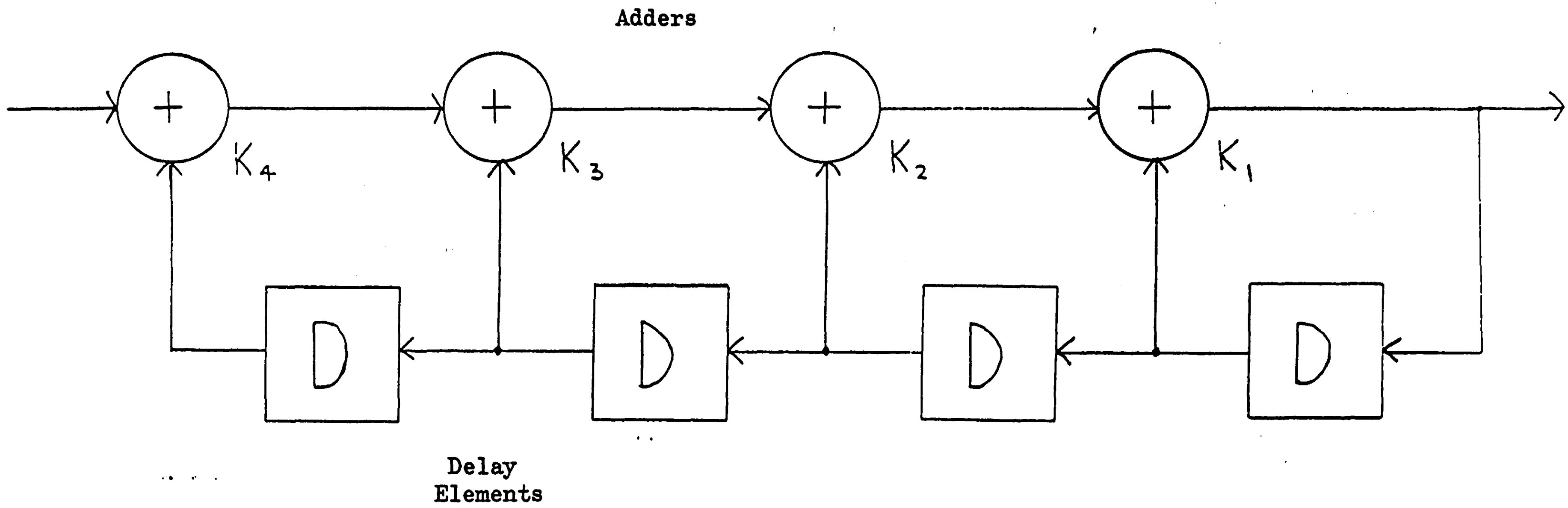

FIGURE 2.5 **Circuit Elements for Linear Prediction**

terms (29). Linear prediction was first applied to speech by Saito and Itakura in 1966 (30) and Atal and Schroeder in 1967 (31) but the linear prediction for a time series had previously been described by Wiener in 1949 (32) and was based on Gaussian linear least squares estimation.

The basic linear prediction equation is

$$S(n) = \sum_{i=1}^{m} a_i \, S(n-i)$$

where S(n) is the predicted value of the nth sample based on the previous m samples, and a_i represents the constant coefficients to be determined.

If an actual sample in a time series is subtracted from the predicted value, the result will be an error known as the prediction error. The linear predictor can be considered to be a filter because it separates the predictable component of a signal from the component it cannot predict; it could therefore be used to separate a signal from background noise. The linear predictor can be used to represent the filtering action of the vocal tract. If the output of such a predictor is compared with the speech samples at the input then the result is the driving waveform i.e. the glottal waveform which drives the vocal tract. This is illustrated in Figure 2.4.

To synthesise the signal at a subsequent time the signal can be reconstructed using the glottal waveform and the prediction values. Figure 2.5 is a block diagram of the circuitry necessary for performing linear prediction. The circuit must contain

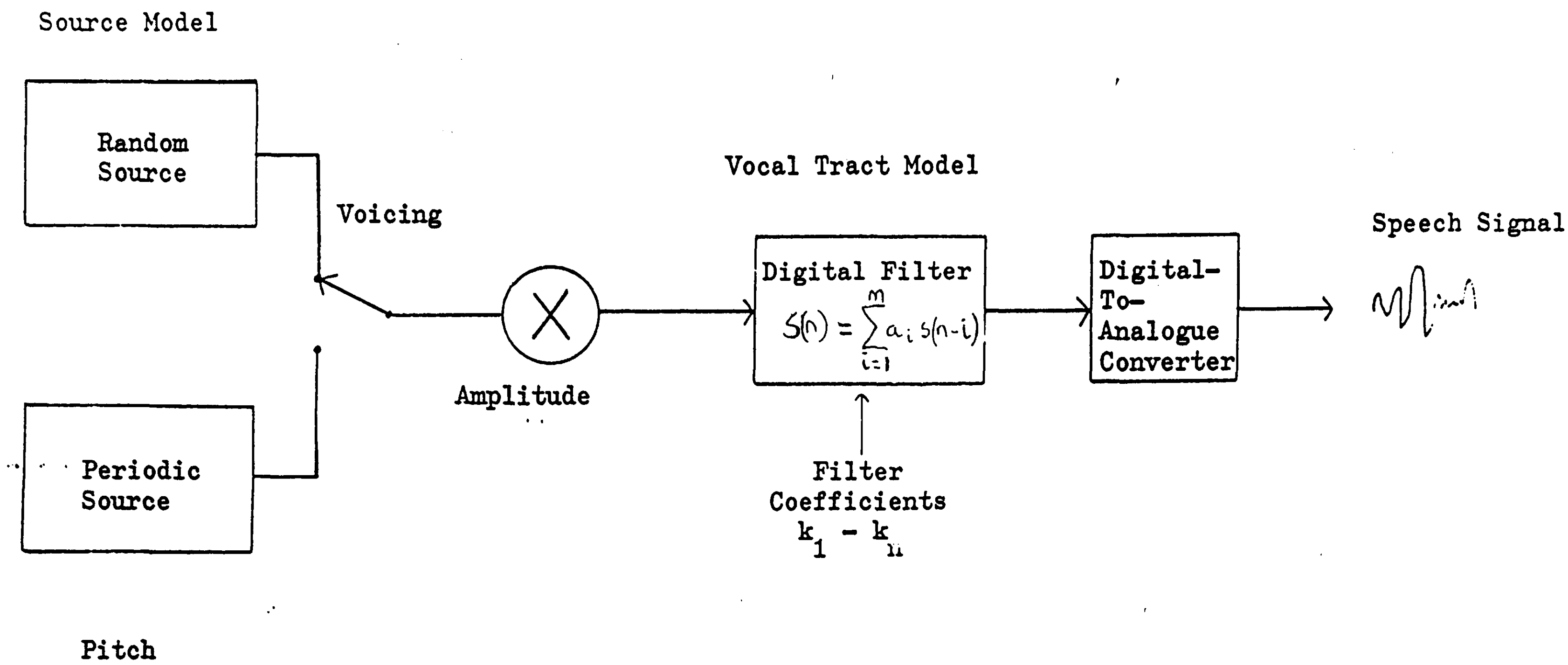

FIGURE 2.6 **The T.I. 5220 Synthesiser**

delay elements for storing previous samples and adders and multipliers to perform the calculations. The prediction coefficients are chosen so as to minimise the prediction error. The coefficients are chosen by solving a set of simultaneous linear equations.

2.4.2 LPC in Practice

Texas Instruments have developed a single chip speech synthesiser to reconstruct speech from LPC data. Figure 2.6 shows the essential features of the 5220 speech synthesiser.

The 5220 chip requires LPC data in the form of a set of coefficients. In this case 12 coefficients are required: energy, pitch, and 10 reflection coefficients which are mathematically related to the prediction coefficients already described. The digital filter requires 20 multipliers and adders to work simultaneously in order to output one speech sample every 125 usec which corresponds to a 8KHz sampling rate. A multiplication and addition must be carried out every 6.25 usec. The 5220 chip does not have the processing power to perform arithmetic at this speed. Instead, parameters are supplied only once every 20msec and the data is modified internally by interpolation.

If a 20 msec frame interval is used with a data resolution of 10 bits per parameter then 6,000 bits would be required for 1 second of speech. However this data rate can be reduced by using a smaller number of bits for some of the parameters, by not storing parameters for silent periods and by reducing the number of coefficients for voiceless sounds.

FIGURE 2.7 Texas Instruments Portable Speech Laboratory

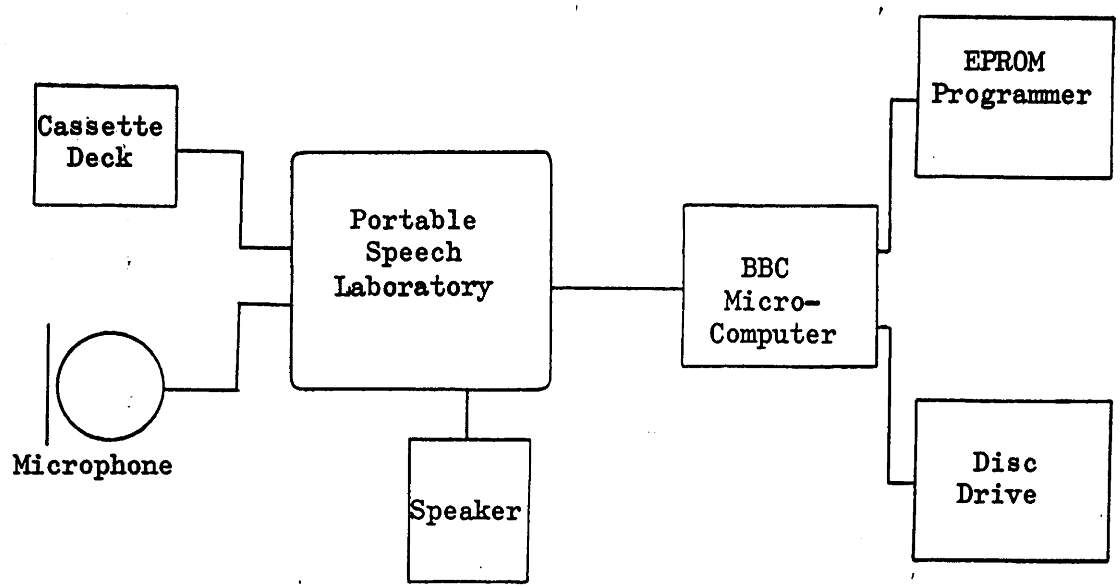

FIGURE 2.8 **The LPC Recording System**

Repeat flags are set if there are no changes in the parameters from the previous frame to further reduce the number of data bits. By using these techniques, LPC can achieve rates as low as 1200 bits per sec.

LPC recording can be of very high quality but there are some limitations imposed on the implementation of the LPC algorithm when synthesis is performed using a single chip such as the 5220 chip. Factors influencing the design of the synthesiser are bit rate, processing power and cost. Texas Instruments have developed a Portable Speech Laboratory (33) for recording LPC data for the 5220 chip and with careful operation and editing of the data good quality recordings can be obtained.

2.4.3 **Recording of LPC Data**

The recording system used in this project was based on the Texas Instruments Portable Speech Laboratory (PSL) (Fig 2.7). A block diagram of the recording system is shown in Figure 2.8. The PSL will analyse a human voice and produce a set of LPC parameters in real time. In recording speech for communication aids it is important to be able to replay the recording immediately through the synthesising system to enable improvements in quality and intonation to be made. The data from the PSL can be edited manually i.e. the energy and pitch values can be altered to desired values for any frame interval. Editing is not a simple task and requires a great deal of operating experience. The PSL displays a list of LPC parameters but does not provide a graphical representation.

Vowels	**Consonants**
uh (the)	p t k
a (bud)	b d g
e (head)	m n ng
i (hid)	
o (hod)	r w l y
u (hood)	
aa (had)	s z
ee (heed)	sh zh
er (heard)	f v
uu (food)	th dh
ar (hard)	ch j
aw (hoard)	h

TABLE 2.1 The Phonemes of British English

Editing is usually restricted to re-recording with a different emphasis by the speaker. When a satisfactory recording of a word or phrase has been achieved the data can be sent directly to the in-built EPROM programmer or it can be transferred to a remote computer over an RS232 standard serial link. A BBC B microcomputer is used to archive speech data. The BBC B is particularly useful for this application because it can also be used as a playback device using its own 5220 speech processor.

2.5 Constructive Synthesis

Constructive synthesis is a technique used for unlimited vocabulary synthesis by concatenating fundamental units of speech. There are many different units of speech that can be used for constructive synthesis. In general, the smaller the unit used the better the quality of the synthetic speech when the units are concatenated. The most commonly used unit is the phoneme and Table 2.1 shows the phonemes of English as given by the International Phonetic Association (IPA). There are 12 vowels and 24 consonants. Phonemes do not take into account the influence of surrounding sounds i.e. the 'k' sound in key is different from the 'k' sound in call due to the influence of the following sounds but they are both classified as the same phoneme. These different 'k' sounds are called allophones of the phoneme 'k'. Allophonic variation in speech occurs because of a mechanism called coarticulation as the articulators move position from one sound to another. There are approximately 128 allophones in standard British English. There are other units which can be used such as syllables or demi-syllables but here the size of the inventory numbers several thousand entries and the control software can be complex. Diphones, units of speech which take into account the transition between adjacent phonemes, offer an attractive alternative to phonemes or allophones. Diphones span from the centre of one phoneme to the centre of the following phoneme. Approximately 2,000 diphones are required for synthesising speech. Phoneme synthesis or allophone synthesis are the most common forms of unlimited vocabulary synthesis. A device using these techniques is the SC01 synthesiser developed by Votrax Inc. The SC01 is a single chip formant synthesiser and is the most widely used constructive synthesis device. The principle used in formant synthesis is to emulate the waveform produced in human speech by copying the pattern of formant frequencies. Formants are resonant frequencies of the vocal tract for a particular sound. The source filter model used in both LPC synthesis and formant synthesis use waveform generators to simulate the sound generated by the larynx and filters to form the frequency characteristics of that sound as it is modified by the articulators. Electronic analog filters are used to simulate the formant frequencies associated with each phoneme. The single chip synthesiser requires a set of codes to specify both the amplitude and pitch of the source signal and the formant frequencies for the analog filters.

The SC01 can be regarded as a sound segment synthesiser as it does not use a strict set of phonemes or allophones but uses a

CONSONANTS WITH MORE THAN ONE PRONUNCIATION

Letter	Pronounced like...	Examples	
c	/K/ or /S/	corn	city
d	/D/ or /T/	paid	passed
g	/G/ or /J/	log	page
q	/K/ or /K+W/	boutique	quit
s	/S/ or /Z/	cats	dogs
x	/K+S/ or /G+Z/	six	exam

LETTER SEQUENCES PRODUCING LONG VOWEL SOUNDS

Long Vowel Sound	Letter Sequence	Example Word
a	ai	maid
a	ay	day
a	a–e	save
e	ee	keep
e	cei	receive
e	ie	retrieve
e	y	every
i	ey	eye
i	y–e	rhyme
i	y	my
i	i–e	line
o	o–e	home
o	oa	goat
o	ow	grow
o	a	war
o	ol	cold
o	or	nor
o	oe	toe
u	oo	soon
u	ou	group
u	u–e	tune

DIPHTHONG LETTER SEQUENCES

Diphthong Sequence	Example Word
ow	power
ou	house
oi	noise
oy	boy
you	you
u	music
ew	few

TABLE 2.2 SC01 Sound Segments

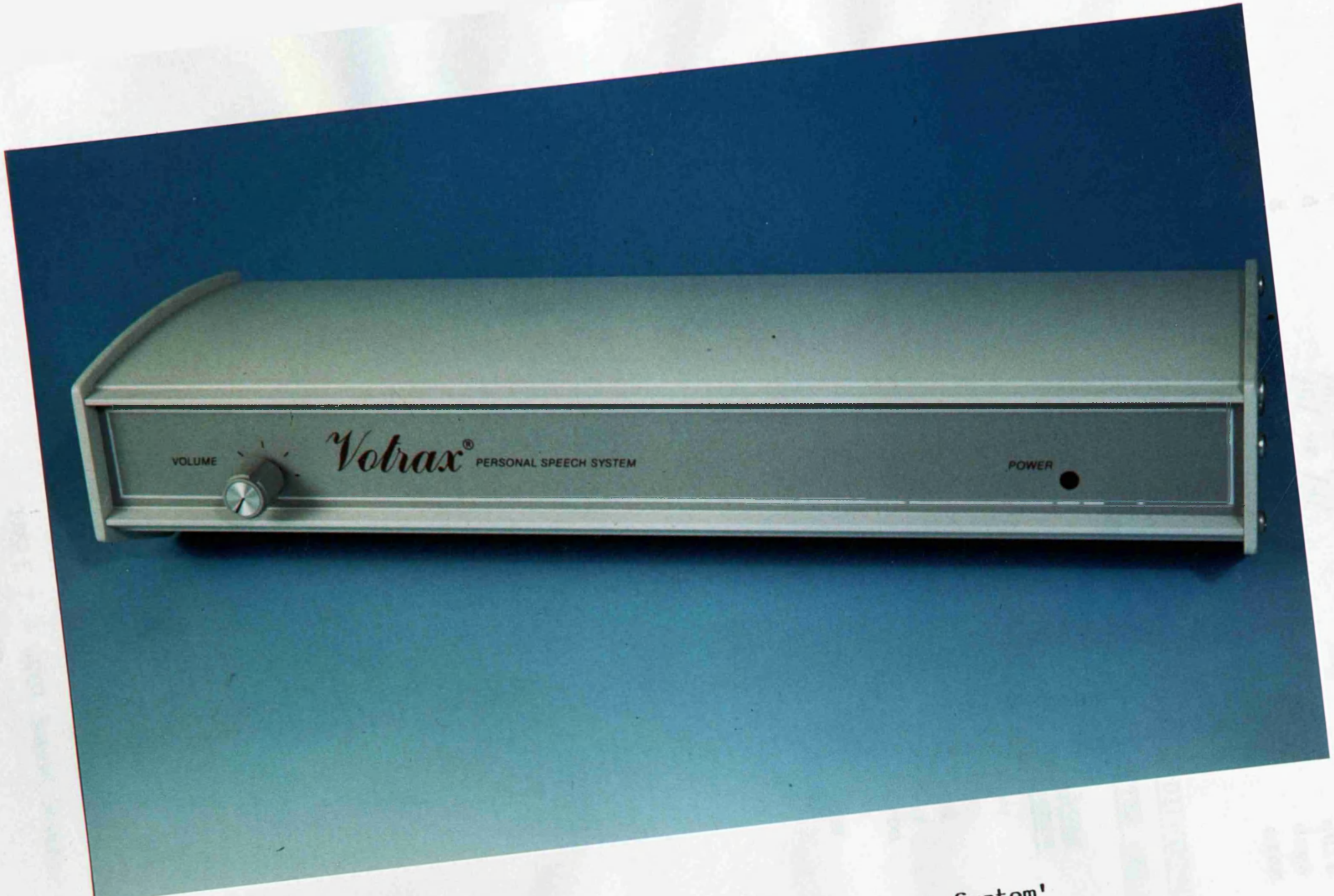

FIGURE 2.9 The Votrax 'Personal Speech System'

SYNTHESIS METHOD	BIT-RATE	QUALITY	RECORDING COST	PLAYBACK COST
Constructive Synthesis	Very Low 100 bits/sec	Low	Very High	Low
Synthesis by Analysis	Low 1200 bits/sec	Moderate	High	Low
Waveform Synthesis	Medium 32 Kbits/sec	High	Low	Moderate

TABLE 2.3 Comparison of the Three Methods of Synthesis

combination of both and has around 80 different sounds as shown in Table 2.2. The Votrax company have also built a text-to-speech module, called the "Personal Speech System", based on the SC-01 chip. A set of rules are used to convert standard text into SC01 sound segments. The device (shown in Figure 2.9) accepts text in the form of standard ascii characters.

2.6 A Comparative Summary of the Speech Synthesis Methods

There are three main features which distinguish the three speech synthesis methods. These features are described below and are summarised in Table 2.3.

a) Bit Rate

Constructive synthesisers store codes related to fundamental units of speech whereas the other methods of synthesis store data for whole words or phrases. The bit rate for constructive synthesis is much lower than synthesis by analysis or waveform synthesis. Waveform synthesis gives a higher bit rate than synthesis by analysis as the stored data is more closely related to the actual digitised sample values.

b) Quality

This feature is examined in Chapter 3. Quality is closely related to the bit rate. In general, the higher the bit rate the higher the quality.

c) Cost

The analysis of human speech to construct a phoneme data base is a complex process requiring high cost computing systems. This is usually done by the manufacturer and incorporated on the synthesis chip. Waveform synthesis (ADPCM) can be achieved for approximately one tenth of the cost of a synthesis by analysis (LPC) method as less computing power is required.

Playback costs are closely related to the bit rate and the size of the stored vocabulary. The cost of the synthesis chips does not vary much but the memory required to store the data makes a waveform synthesis system more expensive than a synthesis by analysis system.

CHAPTER 3 - PERFORMANCE TESTING

3.1 Introduction

There are many features that contribute to the acceptability of a communication aid and some of these were examined in section 1.6. The range of devices developed in this project use speech as the main output mode and it is therefore crucial that the quality of the speech produced is acceptable to both the user and the listeners. This chapter concentrates on measuring the performance of the speech synthesisers. The first trial investigates the intelligibility of speech synthesis devices and the second trial investigates the aesthetic acceptability of the speech produced by the synthesis devices.

3.2 Intelligibility

In order to provide a range of synthetic speech aids several different methods of synthesising speech must be used. A detailed study of the intelligibility of readily available implementations of the techniques used in this project was performed and these techniques were compared to control male and female voices. In order to minimise the variability due to implementation of the various techniques the same sound delivery system was used to play speech to the listeners. This was achieved by recording the speech output from each of the devices on a high quality reel-to-reel tape recorder at the same recording levels. The synthesis techniques have been described in more detail in Chapter 2.

VOICING	NASALITY	SUSTENTION
Veal-Feel	Meat-Beat	Vee-Bee
Bean-Peen	Need-Deed	Sheet-Cheat
Gin-Chin	Mitt-Bit	Vill-Bill
Dint-Tint	Nip-Dip	Thick-Tick
Zoo-Sue	Moot-Boot	Foo-Pooh
Dune-Tune	News-Dues	Shoes-Choose
Vole-Foal	Moan-Bone	Those-Doze
Goat-Coat	Note-Dote	Though-Dough
Zed-Said	Mend-Bend	Then-Den
Dense-Tense	Neck-Deck	Fence-Pence
Vast-Fast	Mad-Bad	Than-Dan
Gaff-Calf	Nab-Dab	Shad-Chad
Vault-Fault	Moss-Boss	Thong-Tong
Daunt-Taunt	Gnaw-Daw	Shaw-Chaw
Jock-Chock	Mom-Bomb	Von-Bon
Bond-Pond	Knock-Dock	Vox-Box

SIBILATION	GRAVENESS	COMPACTNESS
Zee-Thee	Weed-Reed	Yield-Wield
Cheep-Keep	Peak-Teak	Key-Tea
Jilt-Gilt	Bid-Did	Hit-Fit
Sing-Thing	Fin-Thin	Gill-Dill
Juice-Goose	Moon-Noon	Coop-Poop
Chew-Coo	Pool-Tool	You-Rue
Joe-Go	Bowl-Dole	Ghost-Boast
Sole-Thole	Fore-Thor	Show-So
Jest-Guest	Met-Net	Keg-Peg
Chair-Care	Pent-Tent	Yen-Wren
Jab-Gab	Bank-Dank	Gat-Bat
Sank-Thank	Fad-Thad	Shag-Sag
Jaws-Gauze	Fought-Thought	Yawl-Wall
Saw-Thaw	Bong-Dong	Caught-Thought
Jot-Got	Wad-Rod	Hop-Fop
Chop-Cop	Pot-Tot	Got-Dot

TABLE 3.1 Stimulus words used in the DRT

3.2.1 Examples of Different Synthesis Techniques

a) **Phoneme Generation**

The device used was the commercially available VOTRAX Personal Speech System which was the phoneme driven SC-01 speech chip.

b) **Text to Speech (SAM)**

This is an implementation of text to speech on an Apple microcomputer where the speech output is generated from phoneme data via a digital to analogue converter.

c) **Linear Predictive Coding**

The implementation used was the Macleod device (34) which uses the Texas 5220 chip. Both male and female voices were used.

d) **Adaptive Differential Pulse Code Modulation**

The implementation used was described in 2.2.1. Both male and female voices were used.

3.2.2 Measurement of Intelligibility

The speech intelligibility of a device can be measured using the Diagnostic Rhyme Test (35). The DRT measures consonant recognition in the initial position only since these consonants are important discriminatory features for the intelligibility of English words (Table 3.1).

Definitions

For many years linguists and phoneticians have classified phonemes according to features of the articulation process used to generate sounds. These features of speech production are reflected in certain acoustic characteristics which are discriminated by the listener. The following set of distinctive features are used as a basis for classification.

a) **Voicing**

In articulatory terms the vocal chords do not vibrate when the consonants (p,t,k,θ,s,ʃ) are produced, and they do vibrate for (b,d,g,v,ð,z,ʒ,m,n). Acoustically, this means that the voiceless consonants are aperiodic or noisy in character, whereas a periodic component is superimposed on the noise component for voiced consonants.

b) **Nasality**

To articulate 'm' and 'n'
pressure is released through the nose by lowering the soft palate at the back of the mouth. The nasal resonance introduced in this way provides an acoustic difference. In addition 'm' and 'n' seem slightly longer in duration and more intense than their stop or fricative counterparts. The two nasal sounds are the only consonants in this study which lack the aperiodic noise component.

c) **Sustension**

This feature corresponds to the Affrication feature of Miller & Nicely (36). If the articulators close completely, the consonant may be a stop or a nasal, but if they are brought close together and air is forced between them, the result is a kind of turbulence or friction noise that distinguishes (f,θ,s,ʃ,v,o,z,ʒ) from (p,t,k,b,d,g,m,n).

The acoustic turbulence is in contrast to the silence followed by a pop that characterises the stops and to the periodic vowel-like resonance of the nasals.

d) **Sibilation**

This feature corresponds to the Duration feature of Miller & Nicely. The four sibilant sounds (s, ʃ , z, ʒ) are longer and of a higher frequency than the other consonants.

e) **Graveness**

This feature has to do with where in the mouth the major constriction of the vocal passage occurs. Usually, three positions are distinguished. The bilabial and labio-dental consonants (p,b,f,v,m) are grouped as front, the alveolar and dental consonants (t,d,θ,s,ð,z,n) are grouped as middle, and the palatal and velar (k,g,ʃ,z) are grouped as back. The Graveness feature distinguishes front and middle consonants

f) **Compactness**

This is another feature describing the place of articulation; compactness distinguishes back and front/middle consonants.

3.2.3 Experimental Procedure

Ninety-six stimulus words were recorded using a high quality reel-to-reel tape recorder for each of the four techniques. LPC synthesis and ADPCM were recorded using male and female voices.

Twenty-five normal subjects listened to one test tape per day over a period of eight days. The listener's task was to

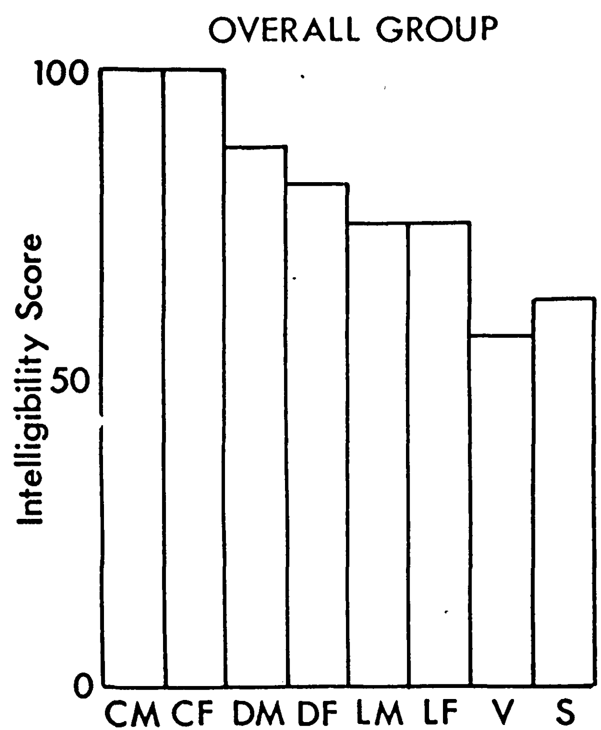

FIGURE 3.1 Intelligibility Scores for the Overall Group of Distinctive Features:
(CM: Control Male; CF: Control Female; DM: ADPCM Male; DF: ADPCM Female; LM: LPC Male; LF: LPC Female; V: Votrax; S: SAM)

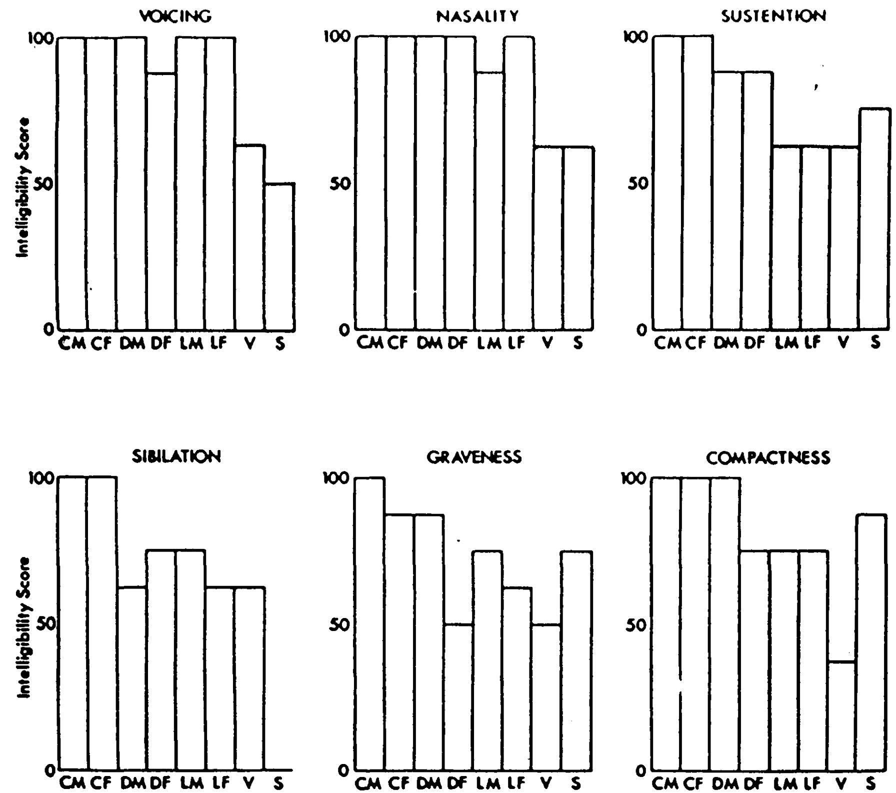

FIGURE 3.2 Intelligibility Scores for the Individual Distinctive Features: (CM: Control Male; CF: Control Female; DM: ADPCM Male; DF: ADPCM Female; LM: LPC Male; LF: LPC Female; V: Votrax; S: SAM)

select the member of the word-pair judged to have been uttered. The tapes were played in a soundproof booth at comfortable listening levels.

The order of presentation of the word pairs was randomised and the actual word spoken for each word-pair randomised for the eight tapes. Each distinctive feature was tested every sixth word-pair. The listener heard one stimulus word every three seconds. The stimulus words used were those of Voiers.

The results are summarised in Figure 3.1 which is a histogram of median intelligibility scores for the overall group and in Figure 3.2 which shows the median intelligibility scores for the individual distinctive features and detailed statistical analysis of this trial is given in Appendix 1a.

Results show that listeners have most difficulty with sibilation in the ADPCM male voice and with graveness in the ADPCM female voice. Both LPC and ADPCM were highly intelligible for voicing and nasality features. LPC speech was significantly poorer than ADPCM speech for sustension. The Votrax performance was poorest for the features of graveness and compactness and SAM was particularly poor for sibilation with an intelligibility score of zero.

3.3.1 Speech Quality

Although synthesis by analysis methods give more intelligible speech than these constructive synthesis methods it is necessary to use constructive synthesis for devices with an unlimited vocabulary.

As it is possible to develop a speech device which combines both methods of synthesising speech, a trial was designed to test the hypothesis that mixing recorded LPC speech with constructed speech is subjectively better than purely constructed speech.

3.3.2 Measurement of Speech Quality

8 test phrases were recorded in four different formats.

A Votrax

B Text to Speech (BBC micro)

C LPC words

D Mixed LPC - Constructed words

The test phrases shown below are from the standard phrase banks used in the Pocket Speech Aids (Chapter 4).

1 HELLO I'M USING A TALKING aid

2 PLEASE GIVE ME A message

3 NO I'D prefer SOMETHING ELSE

4 could YOU phone back LATER

5 PLEASE explain THAT

6 I WANT TO ASK A question

7 CAN YOU explain

8 HOLD ON THIS MACHINE helps

The words shown in lower case are of the constructed type in format D. These words do not occur in the core vocabulary of the 1,000 most frequently occuring English words (37). A core vocabulary of this size can be stored in 2 x 27C512 64Kbyte EPROM devices using LPC synthesis.

Twenty listeners were presented with a test phrase in 2 different formats and asked to decide which phrase he or she preferred. Each listener was asked to judge 48 phrases.

A detailed statistical analysis of this trial is given in Appendix 1a. The main conclusions of this analysis are as follows:

- Method C (LPC words) was preferred to the other methods in almost 90% of the relevant comparisons;
- Method D (mixed LPC and constructed words) was preferred to B (BBC Text to Speech) in almost 90% of cases and preferred to A (Votrax) in almost 70% of cases;
- Method A (Votrax) was preferred to method B (BBC Text to Speech) in 60% of the trials. It was not possible to separate these methods A and B at the 5% level of significance.

This trial indicates ordering for the 4 methods of speech production as:

LPC words > Mixed LPC constructed > (Votrax/BBC Text to Speech.

3.4 Implications for Future Development Work

The first trial shows that LPC speech gives intelligibility scores close to that of ADPCM which requires sixteen times as much memory. The phoneme synthesisers are less intelligible. They, however, offer unlimited vocabulary and lower memory costs.

The Diagnostic Rhyme Test is designed to critically examine the intelligibility of individual words. Phrases are often easier to interpret because the words are in context. The magnitude of the differences in intelligibility score for the various techniques would therefore be expected to decrease if phrases rather than individual words are used.

Intelligibility is affected by signal to noise ratio (38) and there is little doubt that poorly recorded LPC speech, for instance, could be less intelligible than a phoneme based system. In the intelligibility study efforts were made to optimise the recording technique and the settings on each instrument. The intelligibility of the various devices have been assessed in the same form as they are currently being used in low cost portable aids.

In practice, many patients reject devices which use Votrax type speech not purely on account of the limited intelligibility but also because of the unnatural voice quality; the artificial voice causes rejection. The second trial on speech quality examined the influence on acceptability of the degree of naturalness of the synthetic voice.

This trial showed that listeners preferred to listen to the LPC voice than to the phoneme based voices. A significant result from this trial is that listeners would rather listen to phrases which consist of mixed LPC words and constructed words than to purely constructed phrases.

The next three chapters concentrate on the development work which was guided by the results of these performance trials, and also by advice given from experienced Speech Therapists. The speech aids developed can be divided into three types of device. The method of synthesis with the highest speech quality from the trials is used to suit each application.

a) **Direct Selection Devices**

Synthesis by analysis methods give the best speech and ADPCM is used when the memory requirements are low. If the memory constraints are too high then it is neccessary to use LPC synthesis. Chapter 4 investigates the use of LPC and ADPCM in the development of low vocabulary devices.

b) **Coded Devices**

For devices with a vocabulary of several hundred words it is not possible to use the ADPCM method due to the high memory requirements therefore LPC is the method of choice. Chapter 5 describes the development of devices which use a coded input method.

c) **Unlimited Vocabulary Devices**

Unlimited vocabulary devices must make use of constructive synthesis methods. The speech quality trial has shown that text-to-speech on the BBC microcomputer is comparable to the Votrax speech and could therfore be used in preference to the Votrax without loss of quality. Because the overall performance is improved by mixing LPC words with constructed words, a core vocabulary of LPC words with a phoneme back up for words not stored in the memory bank is one possible development for an unlimited vocabulary device. Unlimited vocabulary devices are described in Chapter 6.

CHAPTER 4 DIRECT SELECTION DEVICES

4.1 Introduction

In a major study on the communication interaction between aided and natural speakers sponsored by the International Project on Communication Aids for the Speech impaired (IPCAS) (39), Kraat emphasised the fact that communication devices are a poor substitute for natural speech. Some devices have such a slow communication rate that it is not possible to have an interactive conversation. It is not uncommon to find reported communication rates of as low as 2 words per minute (40). The IPCAS study suggested that a communication aid can not solve all the problems of the speech impaired but can be of benefit in some aspects of communication. The development of a limited vocabulary speech aid capable of real time delivery would help in some situations such as in conversational control or in communication on the telephone. The development of such a device would be of benefit to the group of patients in this study because these patients with acquired speech loss have previously learnt the basic rules for conversational interaction such as the cause and effect nature of communication, attracting attention, holding attention and the rules for turn taking.

This chapter examines the need for a speech aid with a limited number of phrases and describes the design, development and evaluation of such a device.

Aid	Approximate Cost	Manufacturer (M)/ Supplier (S)	Type and Quality of Speech	Input
Votrax Personal Speech System	£375.00	Cyber Robotics, Cambridge (S)	Synthesised speech, Quite good	To ports from computer, etc
Chatterbox	£425.00	QED, Gosport, Hants (S)	Synthesised speech, Very poor	QWERTY keyboard
Convaid	£495.00	Convaid Limited, Eastbourne, Sussex (S)	Recorded speech, Good Pictures and words	Array of 8 x 8 touch switches
Vocaid	£124.00	Texas Instruments (M) QED, Gosport, Hants (S)	Recorded speech, Good	Touch switches 6 x 6 array
Vois 130	£2300.00	P C Werth, London (S)	Synthesised speech, Good	Array of 15 x 8 touch switches
Vois 140	£2300.00	P C Werth, London (S)	Synthesised speech, Good	Numeric key pad + Vocabulary sheets (10)
Possum Speech Unit	£295.00 (+ £825 for Possum)	Possum Controls Limited, Slough, Berks (S)	Synthesised speech, Poor	Via Possum 100SP & Possum switches

TABLE 4.1 Commercial Devices Available at the Start of the Project

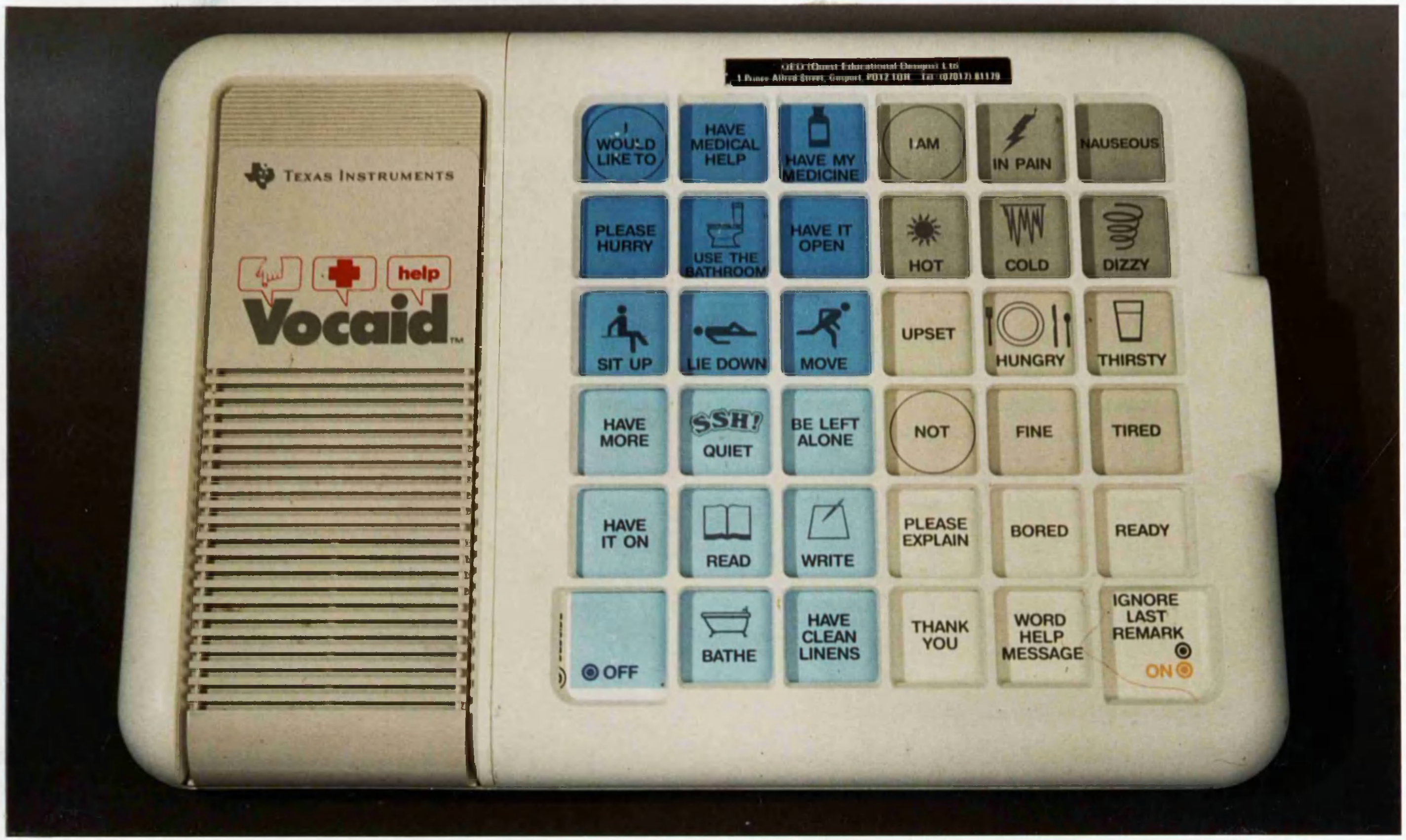

FIGURE 4.1 The Vocaid

4.2 Commercial Devices

A number of communication aids with speech output were available commercially at the start of this project and these devices are summarised in Table 4.1. Of these devices, only the Vocaid (Figure 4.1) can be considered a low vocabulary device. The Vocaid is a direct selection device manufactured by Texas Instruments. The method of synthesis is Linear Predictive Coding and it has an American male voice. Four sets of overlay cards are provided on a matrix of 36 cells, 35 phrases and an on/off cell. Each overlay card is coded and it selects the appropriate vocabulary bank when secured onto the device. Speaking a phrase simply requires the direct selection of one of the cells. Users can make their own overlay cards using Bliss symbols or codes to match the vocabulary provided. One attractive feature of the Vocaid is the price of only £124. This low price has been achieved because the device was originally designed as a mass produced toy for the consumer market and by changing the stored vocabulary, the manufacturer was able to take advantage of the consumer success of the product to provide an aid for the disabled at a very low price. Two disadvantages of the Vocaid for this application were the physical size of the device and its American voice.

4.3 Criteria for a Versatile Low Vocabulary Device

A review of the commercial devices helped to formulate criteria for a low vocabulary portable speech aid. The first and most fundamental point to be considered was conflict between the size of the vocabulary and the speed of operation. The greater the choice of words and phrases, the longer it takes to make the appropriate selection. A device with up to eight phrases can be regarded as capable of producing 'real time' speech. Allowing a greater number of selections than eight places increased demands on the learning power and dexterity of the user. Devices producing 32, 64 or greater number of words or phrases introduce an ever increasing gap between the decision to select a phrase and the final output, whereas an instantaneous response is of a high priority. For less interactive modes of communication, aids with written or visual output are available at lower cost. The primary requirement, therefore was for real time speech.

The next requirement was that the quality of speech should be as good as possible and should be of the correct sex and possibly even of a suitable dialect.

The device had to be simple to operate, battery powered, readily carried and as inconspicuous as possible. A final but important design constraint was that the device should be capable of commercial production for sale at under £200.

4.4 The Development of the Pocket Speech Aid

This device is called the Pocket Speech Aid because of its ability to fit readily into a handbag or man's jacket pocket. It was to become one of the main thrusts of development in the project owing to the encouraging clinical feedback which was generated.

FIGURE 4.2 The Pocket Speech Aid Mk I

4.4.1 The Pocket Speech Aid Mark I

The first version of the Pocket Speech Aid used LPC synthesis and was designed to make use of the TMS 5220 Voice Synthesis Processor's ability to handle serial data and the TI Portable Speech Laboratory facility to store data in this form on a byte wide EPROM device. The eight phrases are stored on one 27C32 4K byte EPROM.

All logic devices are low cost 4000 series CMOS devices and phrase selection and power on/off are simplified by the use of eight double pole/ double throw push button switches. There is no on/off switch on the device because one pole of each selection switch is responsible for applying power to the circuitry and the other pole is used for phrase selection. These features have been implemented to keep power consumption to a minimum and to prolong battery life. The total peak current drawn from the batteries during the speak phase is 55 mA. This gives approximately 10 hours continuous speech from the two PP3 batteries. The device is shown in Figure 4.2 and a full circuit description is given in Appendix 2.

4.4.2 The Pocket Speech Aid Mark II

In order to accelerate the clinical evaluation of the device, a Mark II version was designed. This version incorporated an interchangeable vocabulary cartridge and larger switches which required less activation presure than the switches used in the Mark I device. The circuit board layout was changed but the technology remained the same and the cicuit operation is as in the Mk I device. The interchangeable vocabulary

FIGURE 4.3 The Pocket Speech Aid Mk II

cartridge was introduced to enable the same basic unit to be used by a client in several different situations and helped to speed up the evaluation phase of the device.

Sixteen phrases were stored in an 8 Kbyte 27C64 memory chip. As before, eight switches are used for phrase selection and a ninth switch allows for switching between the two banks of eight phrases. The device is shown in Figure 4.3.

4.4.3 The Pocket Speech Aid Mark III

There were two main reasons for further development of the device.

a) By the second half of the project larger memory devices had become available. The largest EPROM memory chip in 1985 was the 27C512 64 Kbyte EPROM. Although its price at this time was around £50 it was felt that if this price would fall and that an ADPCM version of the Pocket Speech Aid should be developed using two of these memory devices to store the eight phrases. By 1987, the price of these memory devices had indeed dropped to below £20. The ADPCM device was designed using surface mounting technology for the logic components in order to save space on the circuit board.

b) The Speech Quality trial described in Chapter 3 had shown ADPCM speech to be of better quality than LPC and the recording process is much simpler, an important consideration if the device is to be developed as a commercial product.

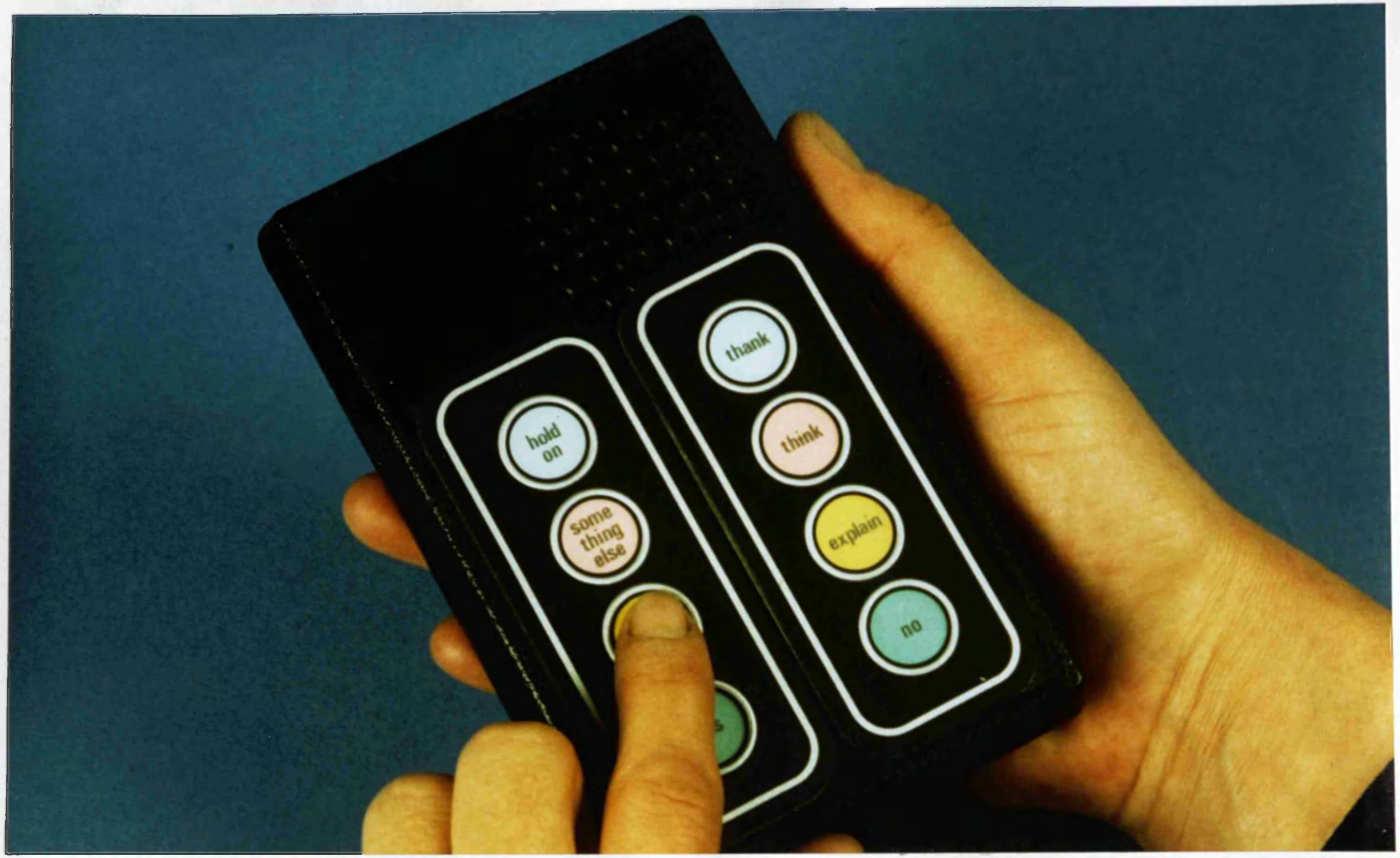

FIGURE 4.4 The Pocket Speech Aid Mk III

As with the LPC versions of the Pocket Speech Aid careful consideration was given to the power consumption of the device. The power consumption is less than in the LPC versions of the device because the ADPCM synthesiser is a CMOS device. Several improvements on the earlier designs were incorporated in the device. Two 4 X 1 low pressure keypads replaced the earlier push button switches as some clients had difficulty in activating the push button switches. Unfortunately, two pole versions of the key pads were not commercially available and as it was still felt desirable not to have an on/off switch careful circuit design was necessary to keep the power consumption as low as possible. The circuit is powered in two stages. In normal stand by mode, the device is monitoring the keypads and the speech chip and audio circuitry is powered down. Only a small current of 10 uA is drawn from the single PP3 battery in stand by mode. When the device is speaking the maximum peak current is around 20 mA. The device is shown in Figure 4.4 and a full circuit description is given in Appendix 3.

4.4.4 <u>The Scanning Pocket Speech Aid</u>

From the initial clinical trials of the Pocket Speech Aid there was a demand from some centres for a version of the device which could be used by patients with Parkinsons' disease. As these patients suffer from hand tremor, a scanning version of the device was therefore designed using the ADPCM method of synthesis and has eight Light

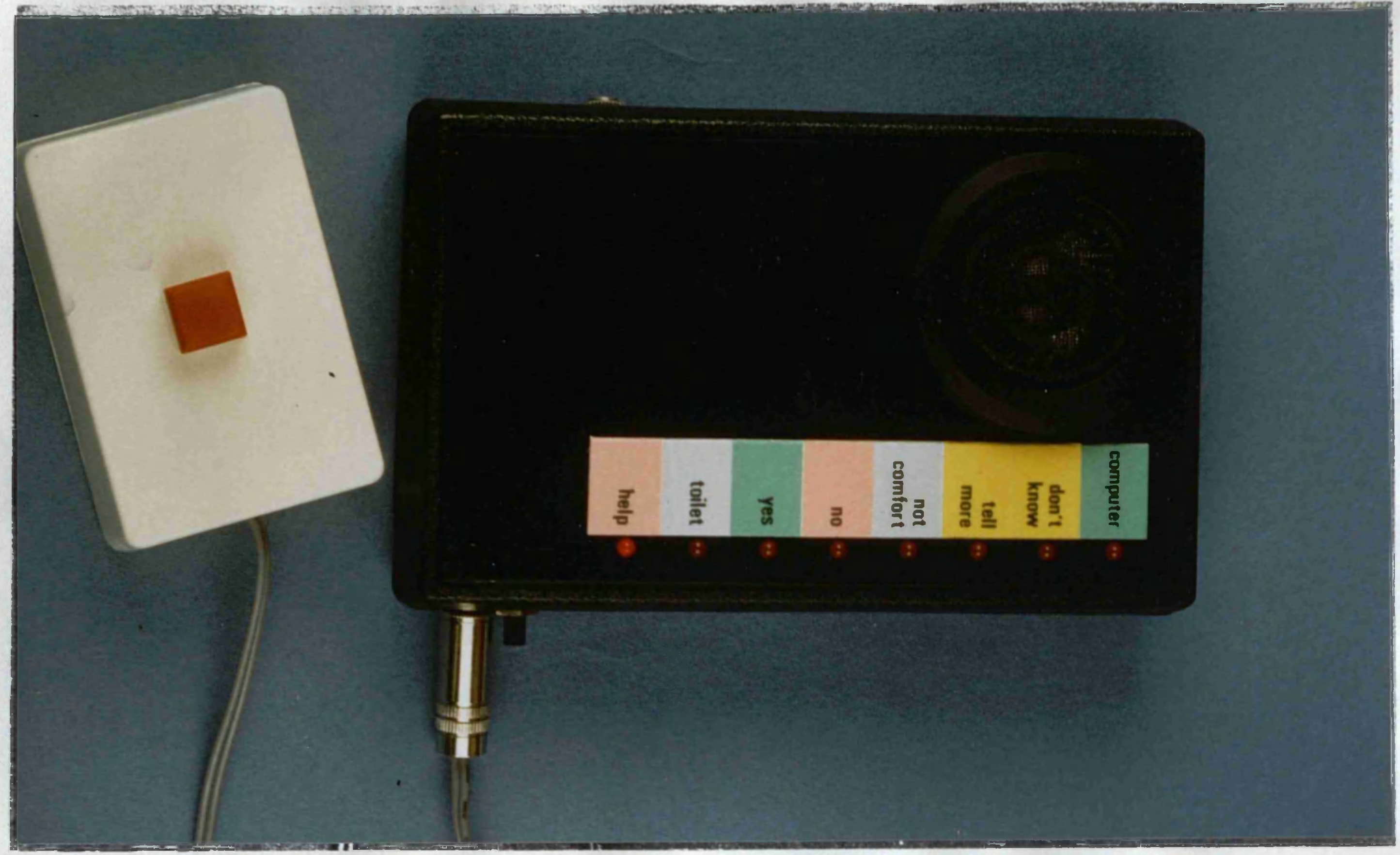

FIGURE 4.5 The Scanning Pocket Speech Aid

Emitting Diodes on the front panel (Fig 4.5). The diodes light in sequence and phrase selection is achieved by activating a single switch when the diode corresponding to the appropriate phrase is lit. A full circuit description is given in Appendix 4. Some applications of the device will now be described.

4.5 Conversation Prompting

Conversation Prompting is a technique of conversational control which can be achieved using the Pocket Speech Aid. The objective is to enable a speech impaired individual to take part in a real time conversation. The use of key phrases such as those stated below help the client to steer the direction of a conversation. The partner does most of the talking but the interest of the client is retained because of the ability to interject and contribute to the communication process.

There is very much more to conversation prompting however than simply using the device. The user must become familiar with the device and must acquire his own particular skills in conversation of this nature. He must make use of every facility available to him such as signs and gestures and not just the Pocket Speech Aid.

4.5.1 Typical Phrases for Conversation Prompting

1 THANKS THATS A GOOD IDEA

2 I'D LIKE TO THINK ABOUT IT

3 SHALL WE TALK ABOUT SOMETHING ELSE

4 I WANT TO ASK A QUESTION

5 NO I'D PREFER SOMETHING ELSE

6 CAN YOU EXPLAIN

7 HOLD ON THIS HELPS

8 YES PLEASE TALK

These phrases have been selected to give the user maximum ability to control a conversation. It is not necessary to speak a complete phrase, the first part of a phrase can be spoken simply by releasing the key at the appropriate time.

The language is limited to eight phrases but the words are not important, only the intentions conveyed in recognised situations. These intentions allow the normal conversational processes of initiation, maintenance and termination.

Initiation - HOLD ON
I WANT TO ASK A QUESTION

Maintenance - YES PLEASE TALK
CAN YOU EXPLAIN

Termination - ID LIKE TO THINK ABOUT IT
THANKS THATS A GOOD IDEA

4.5.2 **Teaching Methods**

The conversation prompting technique has to be taught to a client. Presentation of the device is extremely important. Two instruction booklets have been developed, one to assist the user of a conversation prompter and another to assist people speaking to the user of a conversation prompter. The contents of these booklets are shown in Appendix 5. An audio cassette tape has also been recorded to accompany the teaching booklets. It is hoped that these booklets will help speech therapists in their presentation of the device and help to illustrate the conversation prompting technique to clients.

A video tape of a typical conversation has been made and this helps to illustrate how the client uses the device in conjunction with gesturing and a word board to achieve effective communication. This video presentation is summarised in the following simulated case study.

4.5.3 Conversation Prompting - a Case Study

Margaret is 56 years of age and is in hospital recovering from a recent stroke. She is unable to speak but uses many other forms of communication such as gesturing, facial expression and an alphabet board. Recently she was given a Pocket Speech Aid to help her to gain attention and direct a conversation. The following conversation illustrates how the Pocket Speech Aid can be used to steer the direction of a conversation.

Margaret is sitting by her bed when her 23 year old daughter, Susan, comes in to visit her.

Margaret:	Gestures Hello
Susan:	Hi mum, can you be bothered with a visitor?
Margaret:	YES PLEASE TALK
Susan:	I brought this book for you.

Margaret: THANKS THATS A GOOD IDEA

Susan: How are you feeling since ......

Margaret: HOLD ON....I WANT TO ASK A QUESTION

Margaret has used the speech aid to interrupt her daughter. She then proceeds to use her alphabet board to spell the word 'EXAM'. Susan tells her mother about her recent exam and then talks about her new apartment.

Susan: Our new apartment has a large bathroom and John is thinking about fitting a jacuzzi ......

Margaret: HOLD ON....CAN YOU EXPLAIN

Susan explains what a jacuzzi is and continues to talk about her appartment.

Margaret: HOLD ON....SHALL WE TALK ABOUT SOMETHING ELSE

Margaret interrupts her daughter and gestures for her glasses.

Susan: Do you want me to bring your glasses?

Margaret: NO I'D PREFER SOMETHING ELSE

She uses her alphabet board to spell the word 'DAD` and gestures for her glasses again.

Susan: Oh, you want Dad to bring your glasses tonight?

Margaret: YES HOLD ON

Margaret points to the book.

Susan: Do you want to keep the book?

Margaret: YES

She gestures glasses and points to the book.

Susan: Is it your reading glasses you want?

Margaret: YES THANKS THATS A GOOD IDEA

Although Margaret did not use all eight phrases in the example conversation, she does find all of them useful in helping her to convey different intentions. She does not have to 'speak' the whole phrase but can output part of a phrase simply by releasing the appropriate button at the correct time. It is important to be able to maintain eye contact with your communication partner and Margaret is able to achieve this as she does not have to look at the device when she is operating it. She enjoys using the Pocket Speech Aid and finds it is a useful supplement to her gesturing abilities. When she leaves hospital she intends to use the device as a telephone aid and as a conversation prompter. This will be possible by changing the vocabulary cartridge.

4.5.4 Therapy Sessions

The speech therapist is able to use the instruction booklets and audio and video tapes to demonstrate the use of the Conversation Prompter. The therapist must make the client aware that the device is only to assist the communication process and that other methods must still be used.

Before a conversation takes place there is an intrinsic urge to interact. There is also an identified affinity of interests which might be age, even food, last Saturday's football results or a doctor looking after patients. The situation is then identified e.g. the doctor wants to help this particular patient or the cook wants to know what the client wants to eat and suggests something in response to a reply - "Would you like fish for lunch?". The response could be NO I'D PREFER SOMETHING ELSE or acceptance as in THANKS THATS A GOOD IDEA.

Therapy might be directed towards the exploration and the establishment of these interests and situations, before the spontaneous use of the aid takes place, or indeed while the choice of language for a client is being made.

Modelling practice is often done.

To get attention:	Hold on I want to ask a question
If you want to hear more:	Yes please talk Please explain that
If there is something else to be discussed :	Shall we talk about something else?
To make another suggestion :	No I'd prefer something else

If you are unsure -
The answer might be: I'd like to think about it

But the final use of the aid must be the client's personal application and carry over to their everyday experience or normal interaction. It must be applied to their own personal situations by the client and not imposed by the therapist. The therapist identifies interests, waits for and watches the transition to spontaneous use coming from the client. An imposition by the therapist could in fact impede the existing interaction and the aid could eventually be rejected.

4.5.5 **Evaluation**

Fourteen case studies were undertaken using the Pocket Speech Aid as a conversation prompter (Appendix 6 CS 1-14). Unfortunately the limitation in the number of devices available meant that they often had to be withdrawn from clients who were beginning to use them successfully. It is hoped that commercial availability of the device will soon remedy this situation.

Number of patients	14
Increased communication	10
No increase in communication	4

Seven stroke, two motor neurone disease and a non-vocal patient with unknown etiology increased their communication ability. A motor neurone disease patient and a pseudobulbar palsy patient rejected the device as the preferred mode of communication was writing and gesturing. A motor neurone disease patient rejected the device because he could not accept his prognosis and a Parkinson's disease patient was physically unable to operate the device during the tremor period.

4.6 The Telephone Aid

Although a device specifically designed for telephone communication called Claudius Converse, and manufactured by British Telecom, became commercially available shortly after the start of this project it was decided to pursue use of the Pocket Speech Aid in this situation because of a number of advantages. Whereas the Claudius Converse requires a special telephone socket and has to be plugged into the wall, good quality can be obtained by simply placing the Pocket Speech Aid in close proximity to the telephone handset. The Pocket Speech Aid can therefore be used anywhere and with any telephone system. It is also simpler to use than the Claudius Converse having fewer phrases. Only further experience will tell whether this is a limiting factor or a benefit.

The need for a device for telephone use does not require much elaboration. The knowledge that a device can enable a few selected phrases to be transmitted over the telephone to a friend or relative can be of considerable comfort to a communication impaired client.

Although there can be some overlap between the phrases used in conversation prompting and those for the telephone it would normally be recommended that a few initiating, rather than responsive phrases would be incorporated such as "Please give me a message" or "I need some help". A typical set of eight phrases for telephone use is given below.

1 HELLO THIS IS DAVID KEATING SPEAKING

2 PLEASE GIVE ME A MESSAGE

3 YES THATS FINE

4 NO I'D PREFER SOMETHING ELSE

5 HOLD ON - COULD YOU PHONE BACK LATER

6 I NEED SOME HELP

7 PLEASE EXPLAIN THAT

8 THANKS VERY MUCH - GOODBYE

4.6.1 Teaching Methods

Teaching methods for the Pocket Speech Aid in this mode are not as complex as for the conversation prompter because the use of supplementary skills such as gesturing or writing do not apply in this application. The therapist can still emphasise the use of a set of intentions and explain the use of each phrase as in the conversation prompter.

4.6.2 Evaluation

Ten case studies were undertaken using the Pocket Speech Aid as a telephone device. These case studies are given in Appendix 6, CS 15-24. The results of these case studies showed the Pocket Speech Aid to be of great benefit to most patients and it was therefore decided that the limited number of devices available should be directed towards the evaluation of the device in the more complex conversation prompting mode.

Number of patients	10
Increase in communication	8

No increase in communication 2

Three motor neurone disease, two stroke, a functional aphonia, a pseudobulbar palsy and a multiple sclerosis patient increased their communication ability with this device.

Two patients did not benefit from the device. A stroke patient with right hemiplegia had physical problems in operating the device and a motor neurone disease patient rejected the device because he was unable to accept his prognosis. He had good physical and cognitive skills and could operate the device but was still unwilling to use it as a communication aid.

The Pocket Speech Aid is a very successful device when used as a telephone aid. In this situation it is simple to learn and can help the speech impaired communicate where previously they could not.

Customisation of a few phrases is probably more important in this mode than when the device is used to supplement other communication methods. An example of customisation was given in phrase 1 where the client was provided with the ability to let callers or listeners know who was speaking.

4.7 Special Purpose Use

Banks of phrases have been developed to investigate the use of the Pocket Speech Aid in specific situations such as in a car, at home or for basic needs in a hospital situation. Only a limited number of case studies have been completed for special purpose use because evaluation was concentrated on the conversation prompter and telephone modes of operation. Case Studies 25-29 in Appendix 6 illustrate the possible uses for the device in other situations.

Number of Patients	5
Increase in communication	2
No increase in communication	3

A stroke patient was able to communicate in a hospital ward situation to help her with basic needs. A patient with Pseudobulbar Palsy used the device to discipline his young son. Two stroke patients were unable to use the device in the car situation. One felt sick reading the keys and the other was too confused to use the device. Another stroke patient preferred to write or gesture.

4.8 Summary

The Pocket Speech Aid has been found to be very successful in practice. The response from other centres has been very encouraging and the demand for the device which is currently being manufactured confirm that it has a role to play in assisting some of those with communication impairment. The two most important lessons learned from the trials are that customised recording of phrases is important in some instances and teaching of the device to both the client and the partner requires considerable thought and further development.

4.8.1 Factors Influencing the Acceptance of the Pocket Speech Aid

In addition to the technical features such as speech quality and range of vocabulary which were mentioned in section 4.3, the case studies have shown that there are many other features which influence the acceptance of a device such as the Pocket Speech Aid. Intelligility, sound quality, selection mode, communication rate and range of vocabulary contribute to the acceptability.

a) **Acceptance of Prognosis**

Patients with progressive degenerative diseases often do not accept that their speech or motor functions will not improve with time and there is a tendency to reject an assistive device because of a belief that things will get better soon.

b) **Acceptance of Need**

If the environmental factors are such that a patient relies strongly on the family or carers then there might not be a real need for a speech aid. Other modes of communication such as gesturing or writing might be more efficient.

c) **Timing of Presentation**

If a patient is accustomed to other forms of communication such as gesturing or writing then it is sometimes difficult to introduce a new method. If a patient has been unable to communicate for a long period of time passivity may have been induced.

d) **Identification of Situations**

Speech is a powerful method of communication but a device will be used only if the situation demands its use. There are some situations such as telephone communication where speech is essential and there are other situations where writing or pointing will be preferred.

All these features must be considered by the Speech Therapist when prescribing an aid and when teaching a client its use.

4.8.2 Customised Recording

The final version of the device uses ADPCM synthesis which was implemented to simplify recording and because of availability of high density EPROM devices. A recording station based on a standard microcomputer is being developed at the Department of Clinical Physics and Bio-Engineering. The concept is that major Speech Therapy Departments or Communication Aids Centres could purchase a single recording system which can then be used for supplying the vocabulary for all Pocket Speech Aids in that region.

FIGURE 5.1 The Convaid

FIGURE 5.2 The Vois 140

CHAPTER 5 - CODED DEVICES

5.1 Introduction

The previous chapter has shown that low vocabulary synthetic speech devices can increase the communication potential of many patients. Chapter 6 investigates devices at the other end of the spectrum; unlimited vocabulary devices. Some patients may feel too restricted by a small number of phrases but may not have the physical ability to operate an unlimited vocabulary keyboard device. Devices which use coded input to select a word or phrase offer a vocabulary of several hundred words and these devices could be of benefit to patients with limited finger movement.

5.2 Commercial Devices

A review of commercial devices at the start of this work is given in Table 4.1. The Convaid and the Vois 130 are two examples of devices which are in the medium sized vocabulary range.

The Convaid (Figure 5.1) is a direct selection device which uses the Linear Predictive Coding method of synthesis. The device has 64 cells and each cell is capable of producing 2 phrases, therefore 128 phrases can be spoken using this device. The vocabulary cartridge and word overlay can be changed to provide a new set of phrases. The speech output is an English male voice.

The Vois 140 (Figure 5.2) uses the encoding method for accessing the stored vocabulary. The method of synthesis is phoneme based. The vocabulary size is 1,000 words and word selection is achieved by entering a 3 digit code on the numeric keypad.

The commercial devices available in 1984 did not meet the requirements for the group of patients studied in this project. The Convaid provided good quality speech but the vocabulary could not be easily changed to meet the needs and situations to be investigated. The physical size of the device also limited its use as it was not easily transported for use in remote situations. The Vois 140 provided a large vocabulary but the cost of £2,300 in 1984 was a prohibitive factor. The synthetic speech was phoneme based and from the speech quality results of Chapter 3 it was decided that this was an unacceptable method of producing speech for devices with a limited vocabulary.

5.3 Criteria for a Medium Vocabulary Device

As the commercial devices did not meet the low cost and portability specifications for a speech aid it was decided that a medium vocabulary device should be developed. This speech aid was the first to be developed in this project and it was therefore important that it should be a flexible design for ease of modification to try out different ideas for possible future devices.

It was important firstly to consider the patient skills necessary to operate the device. Patients with good finger movement are likely to be able to operate a keyboard and would therefore be able to access an unlimited vocabulary device. The device was designed for patients with limited finger movement and would therefore use a small number of input switches with low activation pressure. Ease of operation was therefore the first requirement.

The next requirement was that the speech quality should be the best available, of the correct sex and of a suitable dialect. Because the vocabulary size was to be several hundred words ADPCM synthesis could not be used but LPC could be implemented.

5.4 A Voice from the Past

In 1983, Major Norman Macleod, a remarkable patient suffering from motor neurone disease, who realised that his speech would soon be lost, had the foresight to select a 625 word vocabulary and record these words on a cassette recorder. Major Macleod was a highly motivated individual who hoped to have his vocabulary transferred to a synthetic speech aid. A portable low power device accommodating the specific ideas of this individual patient, would meet the criteria of section 5.3 and would be sufficiently versatile to try out different methods of selection of words or phrases. The device is named after this particular patient and is known as the Macleod unit.

5.4.1 Design Features

The practical design features of portability, simplicity of use and good quality speech were added to the technical considerations of low power consumption and versatility. As LPC speech was to be used, this required a relatively large memory, about 1K bytes per six words. The word selection mechanism suggested by the patient required five

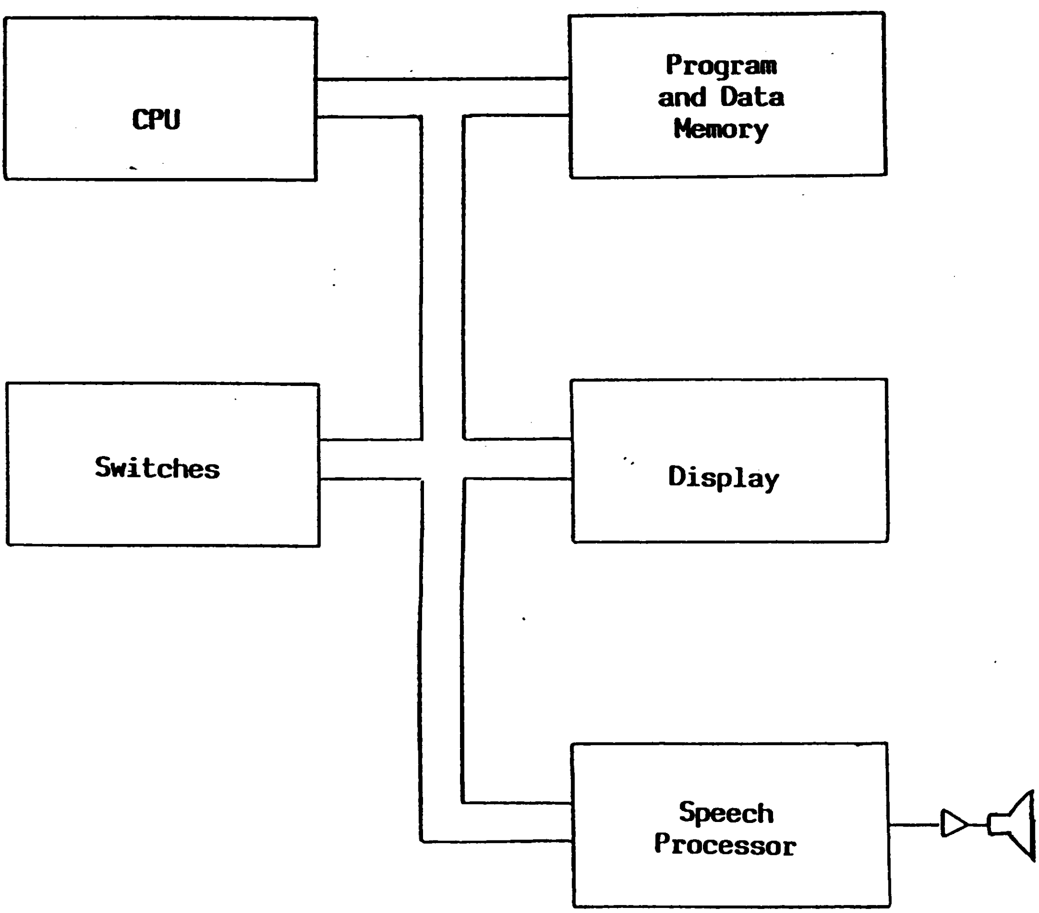

FIGURE 5.3 The Macleod Unit: Schematic

selection switches, together with SPEAK and CLEAR switches. Each word was selected using a four-digit numerical code, each digit being to the base five (for example 4032). To assist the speaker, an alphanumeric liquid crystal display was incorporated which displayed the code and the selected word. In general design terms, the seven switches could be mounted on the front of the device and could be reasonably spaced so that only a modest amount of finger control was necessary for operation.

Finally the device was to be portable - and therefore battery powered - and should incorporate automatic power saving circuitry.

5.4.2 Hardware Implementation

Having specified LPC speech, the Texas Instruments 5220 speech chip was used. The code selecting switches allowed for the 625 words which had to be selected by the patient, but it was decided to limit the device to 64K of memory (60K of speech data) to avoid page selection in the memory addressing. At the time, this was partly an economic decision because of the high cost of CMOS memory. 8K byte memory devices were the largest available memory devices.

The NSC800 (National Semiconductor) was used as the processor because of its low power consumption, its power save features and the availability of suitable support chips.

The final system (Figure 5.3) incorporates 62K of CMOS ROM, 2K of RAM, four byte wide I/O ports and two 4 bit ports. Two ports control the speech chip, two control the display, one

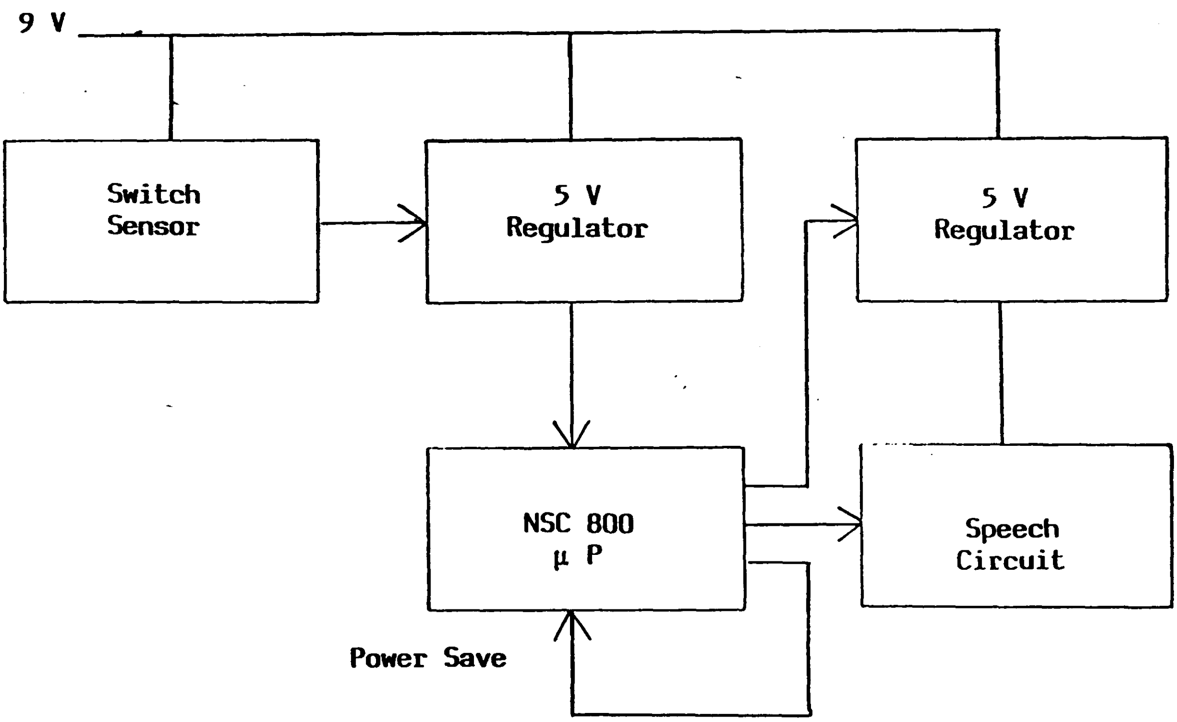

FIGURE 5.4 The Macleod Unit: Power Save Circuitry

Mode of Operation	+ve 9 V Supply	−ve 9 V Supply
1. Switch Monitoring	0.05	0.05
2. Processor Power-Save	10	0.05
3. Processor Running	15	0.05
4. Speech Output	50	30

TABLE 5.1 Macleod Unit: Current Consumption (in mA) in the four modes of operation

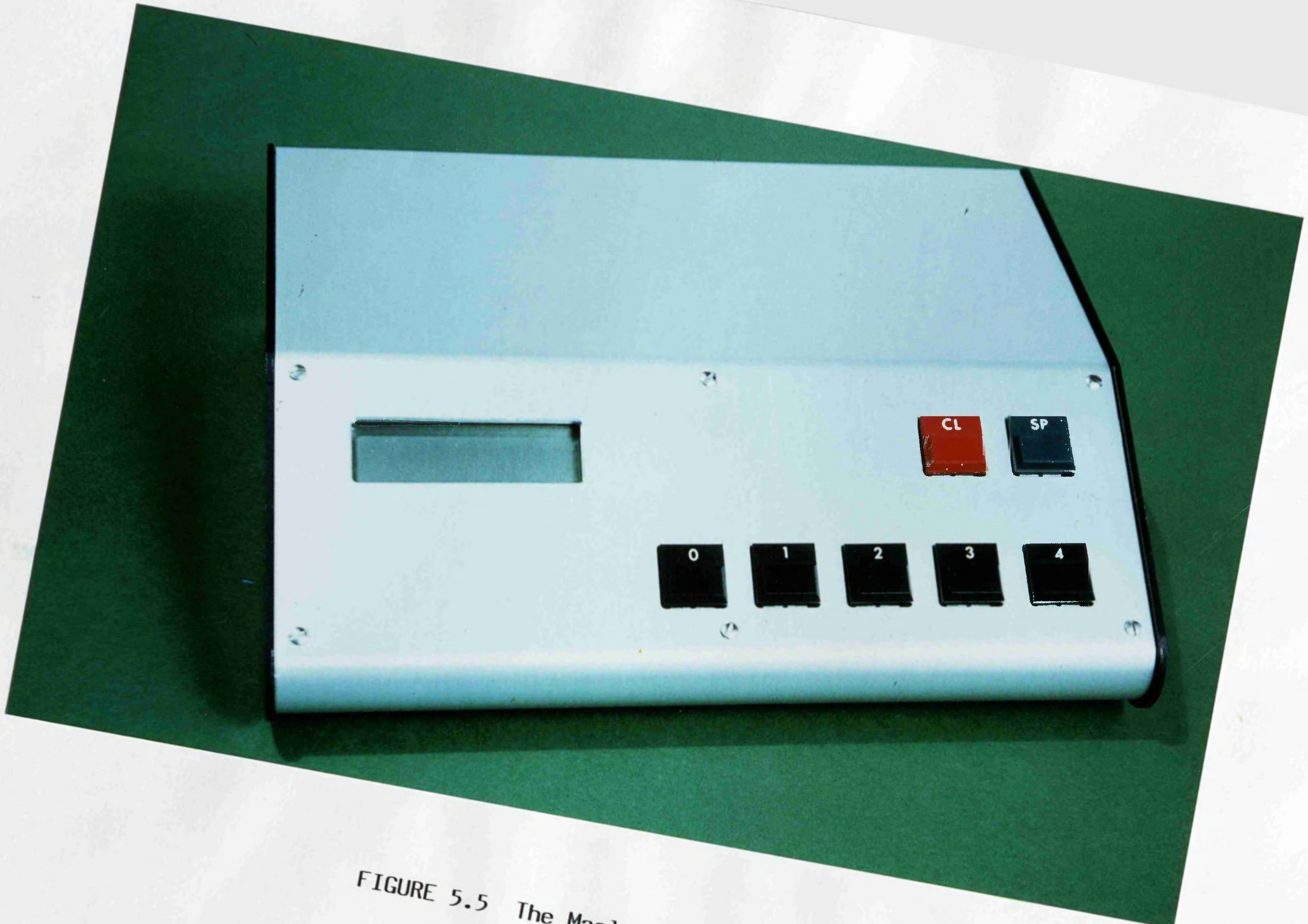

FIGURE 5.5 The Macleod Unit

monitors the switches and the final port is free for further expansion or modification. When speech is produced at the loudspeaker the current consumption is at its maximum, 50 and 30 milliamperes being drawn from the two 9V (PP3) batteries. These currents are neccessary to maintain the quality of the speech output, but fortunately they are required only for relatively short times, that is, when the device is 'speaking`. Most of the time the device is waiting further instructions from the user and so considerable effort has been expended in minimising the current consumption in this mode; three levels of power down are provided (Figure 5.4). Three CMOS integrated circuits are powered continuously from one of the batteries, these circuits monitor the switches and require a current of only a fraction of a milliampere. When a switch is pressed, power is applied to the processor and, after the reset time of 600 msec, the program starts and the switch port is read. While waiting for the next switch to be pressed the processor enters its power-save mode. When 'SPEAK' is pressed, and the speech processor is presenting a signal to the audio amplifier, the current consumption is at its greatest (table 5.1). Finally, after 2 minutes of inactivity, the power to the processor is switched off and the device returns to its initial state of just monitoring the switches. Battery life is very dependent on both the total time of use and the frequency of use, but is of the order of weeks. The device is shown in Figure 5.5 and a full circuit description is given in Appendix 8.

5.4.3 Software Design

The control program consists of a suite of subroutines written in Z80 assembly language in order to use as little memory as possible. Four sequential switch selections are read and converted to a number which determines the location of the speech data by using a jump table.

The data consists of the ASCII letters of the word for the display, two bytes denoting the length of the LPC data and then the LPC data itself. Some care had to be taken to ensure that the speech processor operated reliably when used in combination with the power-up, power-down sequences. Briefly, after power-up of the speech processor, a 500 msec delay was allowed before resetting the device with a sequence of nine bytes of all ones (OFFH).

5.4.4 **Selection of Vocabulary**

The vocabulary for Major Macleod was his own choice and the words were arranged in subject groups at his request. The selected words and phrases were specific to his own needs and situations and would not be applicable to any other patient's requirements. A more general core vocabulary was required.

The frequencies of word occurence in communication samples produced by adult communication aid users were examined by Beukelman et al (41). The 500 most frequently occuring words represented 80% of the total words in a combined sample of 5 communication aid users. However of all the messages generated by the subjects only 33% could be communicated in their entirity using words from the 500 most frequently occurring words. This is because grammatical or structural type words are much more frequent than content words.

Because users of a limited vocabulary device could not be expected to generate complete syntactically correct sentences, content words were preferred to grammatical words. The core vocabulary is investigated in more detail in chapter 6 but for this application word frequency was used only as a guide for selecting a general vocabulary and greater emphasis was placed on the content words necessary for basic communication in limited situations. For example, if the device was to be used in a home or hospital situation the basic needs vocabulary developed by Shane (42) would be incorporated.

5.4.5 Method of Use

In common with the other commercial devices in this category, the Macleod unit is used in conjunction with a large table of words. The words are arranged in groups, for example; verbs, clothing, food, etc. This arrangement certainly assists the initial presentation to the user who quickly learns to locate the desired word. Each word is accompanied on the table by its four-digit code (for example 4032). Message or sentence selection is accomplished by entering sequences of the four digit codes into the device by means of these numbered switches. Each selected digit is confirmed on the LCD display and each selected word is also confirmed when the fourth digit is pressed. Any word selection error can be cancelled using the 'CLEAR' switch. Pressing the 'CLEAR' switch a second time erases the whole buffer in preparation for a new message.

When 'SPEAK' is pressed, the words which have been selected to form the message are spoken and also appear in sequence on the display. Pressing 'SPEAK' again repeats the message. Sentences of up to 128 words can be composed and spoken. The user is restricted to choosing from the number of words initially placed in the device's memory (typically 400). This restriction is common to devices using the LPC method of synthesis but the quality of the speech output is better than that of the simple constructive synthesis device. By retaining a core of commonly used words, customisation of each unit to individuals can be limited to providing a few relevant names and phrases. This customisation has to be performed using the Texas Instruments Portable Speech Laboratory.

The first prototype device was completed using the Major Macleod's own voice digitised and stored in the memory. This prototype allowed the confirmation of the design and gave the patient a communication aid fairly quickly when he began to need it most. Because his voice was beginning to deteriorate when the recording was made, the quality of some of the speech proved inadequate in practice. In a later version of the vocabulary the initial recordings were supplemented with another selected speaker and the quality of the speech output was improved.

5.4.6 **Evaluation**

Number of patients	4
Increase in communication	2
No increase in communication	2

Four patients have used the Mcleod unit as a communication aid. Communication Aid case studies are given in Appendix 6 for these patients (CS30- CS33).

The vocabulary size was between 400 - 600 words and was customised for each individual patient. The patients were provided with a 'voice' of the appropriate sex and accent. Words were grouped under different headings for all four patients. The code chart was re-designed using colour coding and alphabetical ordering of the words for patient CS31. The device was used by all four patients in the home environment.

A motor neurone disease patient and a stroke patient increased their communication ability using this device. The motor neurone disease patient was involved in the design of the original specifications for the device and was highly motivated to use his own custom built speech aid. The stroke patient regained his desire to communicate when using the device, eventually using his own speech.

Two motor neurone disease patients found the device too slow to operate and both were frustrated by the limited vocabulary.

5.4.7 **Summary**

The clinical feedback showed this device to be useful for Major Macleod who was highly motivated to use a speech aid designed to meet his own specifications. It must be emphasised that Major Macleod was not a typical example of a communication aid user and the device was much less successful with other patients. The communication rate was a difficult problem. The patients experienced frustration in searching for words in the code chart, often finding the word they were looking for was not in the vocabulary.

In general, the device has limited use by the acquired disability group. These patients are suffering from progressive disorders and their morale is often low. The cognitive effort to select codes is too great for them and usually outweighs their motivation to use the device.

Some patients suffering from congenital speech disorders are progressing as far as cognitive skills are concerned and these patients can be highly motivated. For example, one spastic child at the Corseford School for Spastics has memorised over 500 codes and would therefore be able to use a device such as the Macleod unit.

The remit for this project is for patients with acquired disability but further investigations with the congenital speech disorder groups will take place under the guidance of the staff at Corseford School.

6.1 Introduction

A device capable of producing any desired word or phrase would be a useful addition to the range of synthetic speech aids developed for this patient group. Unlimited vocbulary devices use a keyboard as the input mechanism and are considered to be direct selection devices because each letter is associated with one particular key. A major advantage of the direct selection approach is that it is straightforward and no learning of a code is required.

A major limitation of keyboard operated devices is that they require a greater range of motion control than devices with simpler input mechanisms. Because good physical skills are required to operate such a device only a small number of patients in the acquired speech loss group would be able to use it. One particular group of patients who could benefit from an unlimited vocabulary device are those patients who are still in the early stages of a progressive disorder and whose disease has not progressed far enough to affect finger movement.

The message preparation time would make it unlikely that a keyboard operated device could be used to achieve an interactive conversation (43). If messages could be prepared in advance by the patients themselves then phrases could be constructed to suit a particular situation. This would help to increase the communication rate of an unlimited vocabulary device. In this way, the limitations imposed by the limited vocabulary device such as the Macleod unit would be removed and the positive features of the Pocket Speech Aid could be incorporated.

This chapter examines the design, development and evaluation of unlimited vocabulary speech aids.

6.2 Commercial Devices

A number of commercial speech synthesisers offer an unlimited vocabulary. Although expensive systems such as DECTALK are available which produce good quality synthetic speech, the computing power in the lower cost devices which are based on microcomputer technology produce speech of much poorer quality.

Low cost devices use phonemes or allophones (44) as the fundamental unit of speech.

The devices which were available at the start of this work are summarised in table 4.1. The Chatterbox requires phonetic input and is therefore more cognitively demanding to operate than a text input device. A device which can accept text input is the Votrax Personal Speech System. This is the most widely used phonetic synthesiser and it can be driven by a keyboard or any device capable of producing ASCII characters in a standard serial or parallel format (Fig 2.9). The device incorporates a text to speech algorithm and standard English text will therefore be converted into speech thus making the device cognitively less demanding than phonetic input devices.

FIGURE 6.1 Text-to-Speech Mk I

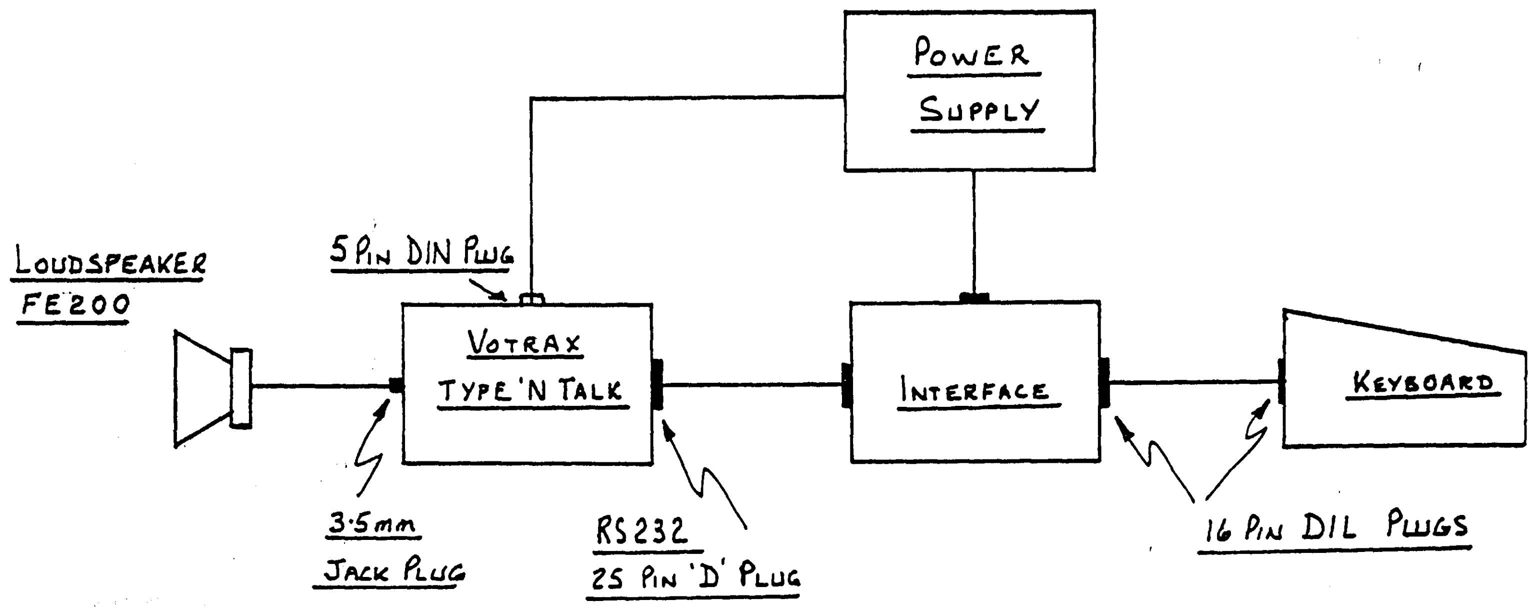

FIGURE 6.2 Text-to-Speech Mk I: Schematic

6.3 Text to Speech Mark I

The first unlimited vocabulary device developed in this project was for a female motor neurone disease patient whose occupation was that of a clerkess/typist. She was unable to speak intelligibly but remained a fluent typist.

It was decided to make use of the commercially available Votrax Personal Speech System and to interface the Votrax with a standard ASCII keyboard. A circuit board was designed to provide intelligence to the system. The use of a microprocessor and non-volatile memory allowed a repeat facility and phrase storage to be incorporated into the unit. The keyboard, interface card, power supply and the Votrax were packaged in one convenient enclosure for ease of transportation. The unit is shown in Figure 6.1.

6.3.1 Design Features

The Votrax Type 'n' Talk is not based on CMOS technology therefore it is not possible for this device to be battery powered but it was decided that it should be portable enough to allow for easy transportation.

The Votrax device will convert text represented by ASCII characters sent as parallel or serial data into synthetic speech. The speech quality of the Votrax was examined in detail in Chapter 3.

6.3.2 Hardware Implementation

A schematic diagram of the system is shown in Fig 6.2. The interface board was based on the Z8671 microcomputer chip which is a Z8 microprocessor with on board BASIC interpreter, serial and parallel ports. This particular device was chosen to reduce the number of support chips neccessary and to simplify the interface board. The device has software selection of baud rate and has an auto-start feature for EPROM stored programs.

6.3.3 Software Design

The control program is written in the Z8 interpreter language 'Tiny Basic' which is a subset of the Microsoft industry standard. The program is responsible for monitoring the keyboard to determine if the keys pressed for phrase storage or recall have been entered. If a phrase is to be stored then the text string is stored in non-volatile RAM and if a phrase is to be recalled then the text is sent to the Votrax for output. In normal text to speech operation the text is sent to the Votrax via a buffer area of memory.

6.3.4 Method of Use

Text is entered from the keyboard and is automatically stored in the non-volatile RAM until a carriage return <CR> is entered. A further <CR> causes stored text to be passed via the RS232 serial link to the Votrax. Further entries of <CR> causes text to be repeated.

The interface can store 24 (64 byte) phrases which may be recalled by a simple two character code. Storage of a phrase is achieved by typing a "*" followed by a letter code "A to X". This is followed by the required text terminated by <CR>. Any previously entered phrase can be recalled by typing "£" and the letter code "A to X". The text stored for that phrase is passed to the Votrax for synthesis by typing <CR>.

A detailed hardware and software description is given in Appendix 8.

6.3.5 Evaluation

Five case studies were undertaken using the Votrax Text to Speech device. These case studies are given in Appendix 6, CS 34-38.

Number of patients	5
Increase in communication	0
No increase in communication	5

Two motor neurone disease, a multiple sclerosis, a pseudobulbar Palsy and a stroke patient rejected the device as a communication aid. The case studies show that the main reason for rejection of the Votrax as a communication aid is the mechanical sounding voice. The results of these clinical trials prompted the speech quality trial described in Chapter 3.

One patient (CS 34) used the device to communicate with her grandchildren because they enjoyed hearing the robotic voice and were unable to read her written notes which, at the time, were her preferred mode of communication. The clinical trials have shown the Votrax device to be an unacceptable communication aid for patients with acquired speech loss. The device may be more acceptable to patients with congenital speech

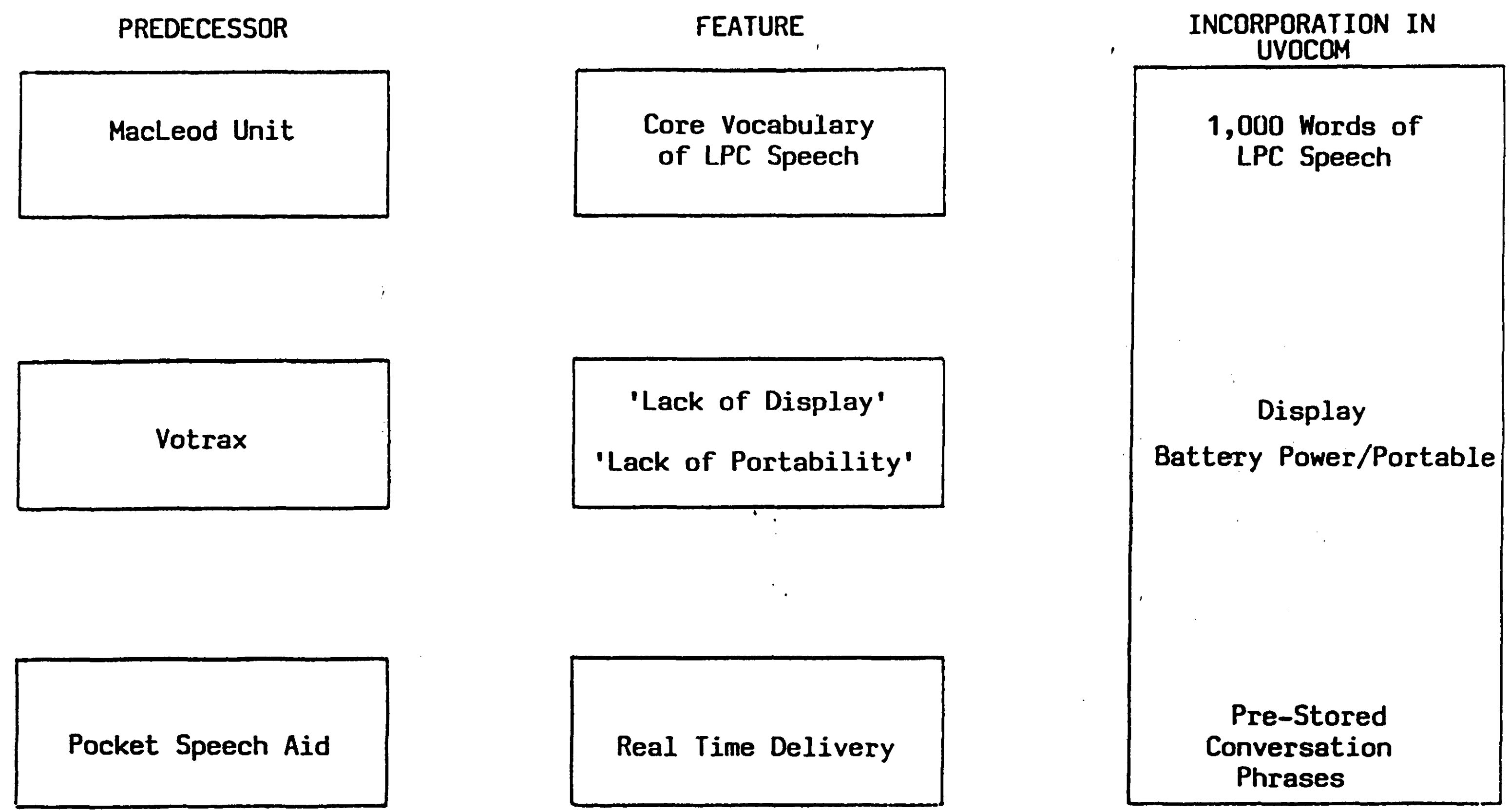

FIGURE 6.3 Evolution of the Uvocom

disorders because these patients have not been used to producing natural sounding speech and would not be embarrassed by the mechanical sounding voice.

The experience gained from the Votrax device and the results of the speech quality trial indicate that an unlimited vocabulary device with speech quality closer to the synthesis by analysis methods would be more acceptable to clients and could be of considerable value.

6.4 The Evolution of the Uvocom

The experience gained from the clinical trials of the Votrax, Macleod unit and the Pocket Speech Aid contributed to the design of an improved unlimited vocabulary device called the Uvocom (Unlimited VOcabulary COMmunicator). Salient features from each of these devices were incorporated in the Uvocom and this is illustrated in Figure 6.3.

In Chapter 3 it was shown that listeners prefer to hear speech generated from synthesis by analysis methods rather than that produced from constuctive synthesis methods. Case studies have confimed this fact.

Trials with the Votrax device indicated that the main cause for rejection of the device was the unnatural sounding speech. The nature of the synthesis by analysis method means that every word in the vocabulary must first of all be recorded. It is therefore not possible to develop a portable and reasonably priced unlimited vocabulary device using only this method. A hybrid device which has a large vocabulary of pre-recorded words with phoneme back up would be more acceptable than a simple constructive synthesis device. Chapter 3 has shown that listeners prefer to listen to phrases constructed in this way to phrases constructed entirely from constructive synthesis. The Macleod unit used a vocabulary of pre-recorded LPC words which was of acceptable quality to most patients and this would be extended in the Uvocom.

The Votrax device highlighted the need for the device to be portable. One disadvantage with the Votrax was the requirement for a mains power supply. To overcome these limitations, the Uvocom should be battery powered and portable enough to fit into a bag or brief-case. Users of the Votrax also commented on the need for a display to check pre-stored phrases before speaking them. Double key presses could not be detected and this affected the quality of the spoken word because the entered text string is converted to phonemes to drive the synthesiser. A display would therefore be incorporated in the Uvocom.

The Pocket Speech Aid had been the most successful device of the project and the key feature of this device was its real time delivery. The Uvocom would therefore incorporate a set of pre-recorded conversation phrases to enable interactive communication to be added to the unlimited vocabulary feature. The Pocket Speech Aid had only a limited number of phrases to facilitate selection. In the Uvocom the interactive phrase keys would be separate from the main keyboard and therefore easily found.

6.4.1 The Core Vocabulary

There have been many studies of the frequency of English words but most are based on written English. Two published word lists are based on spoken English and these are the word lists used in the core vocabulary of the Unlimited Vocabulary Communicator (UVOCOM) (37, 45). Many authors of these word lists make the point that a very small number of very frequently used words make up a considerable proportion of the language produced. Schonell (45) claims that 10 words account for 25% of spoken English. Grammatical or structural words are very much more frequent than content words, the ten most frequent being I, the, and, to, a, of, be, in, we and have. Although these words have a very high frequency of occurence they would be of little use in phrase construction because no content words are included. If the core vocabulary is increased to 50 words then this would account for 50% of spoken English. The first 1,000 most frequently occuring words account for 94% of spoken English and a higher proportion of content words are obviously included in this list.

The thousand most frequently occuring words may not be the optimum core vocabulary for a communication aid user because the published word lists are based on normal adult speech and not on that of communication aid users. Comprehensive lists for communication aid users have not been published.

In 1986, 1,000 LPC words could be stored in 4 of the largest available memory chips (27C512 Eprom). It was therefore decided to develop a device based on this size of core vocabulary.

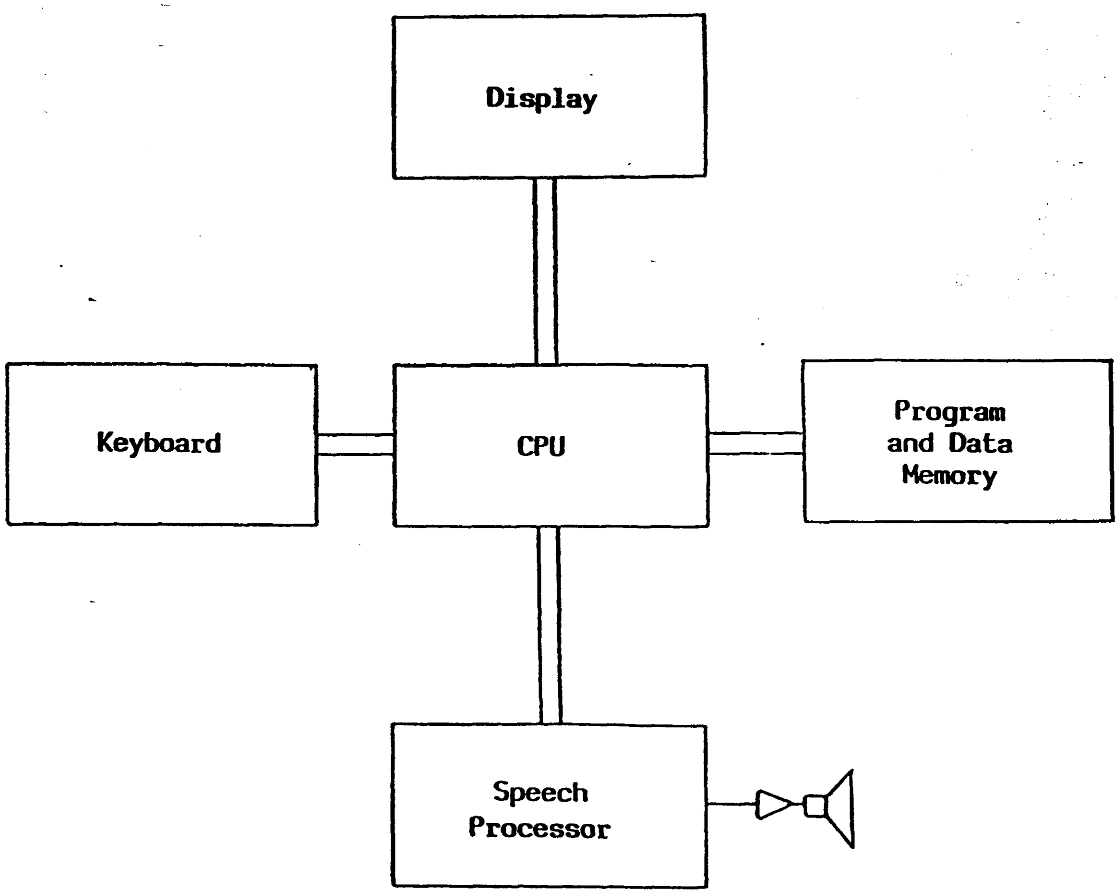

FIGURE 6.4 The Uvocom: Schematic

6.4.2 Hardware Implementation

The use of commercially available modules would reduce the development time for the Uvocom. A useful piece of software had been produced for the BBC microcomputer model B and would convert phoneme codes into LPC code to drive the 5220 synthesiser. This software is avalable on EPROM and is written in 6502 code so that a CMOS version of the 6502 processor would be a suitable microprocessor for the Uvocom.

The decision to use LPC speech for the core vocabulary dictated the use of the Texas 5220 speech processor. The core vocabulary of 1,000 words is stored in four 64 Kbyte EPROM devices (27C512): the Uvocom would be a CMOS version of the 6502 processor.

The final system (Fig 6.4) incorporates 256 Kbytes of CMOS memory of which one 64 Kbyte EPROM is reserved for Program memory, 32Kbytes of RAM and 2 byte wide input/output ports which control the QWERTY keyboard, speech processor and printer output.

A full description of the Uvocom hardware is given in Appendix 9.

6.4.3 **Software Design**

The BBC model B microcomputer was used as a development tool for the Uvocom device because it uses the 5220 speech processor and also because firmware was available for Phoneme - LPC conversion. Chapter 3 compared the quality of the constructed speech formed by this method with the Votrax device and there was no significant difference in speech quality. The software to perform text to phoneme conversion was

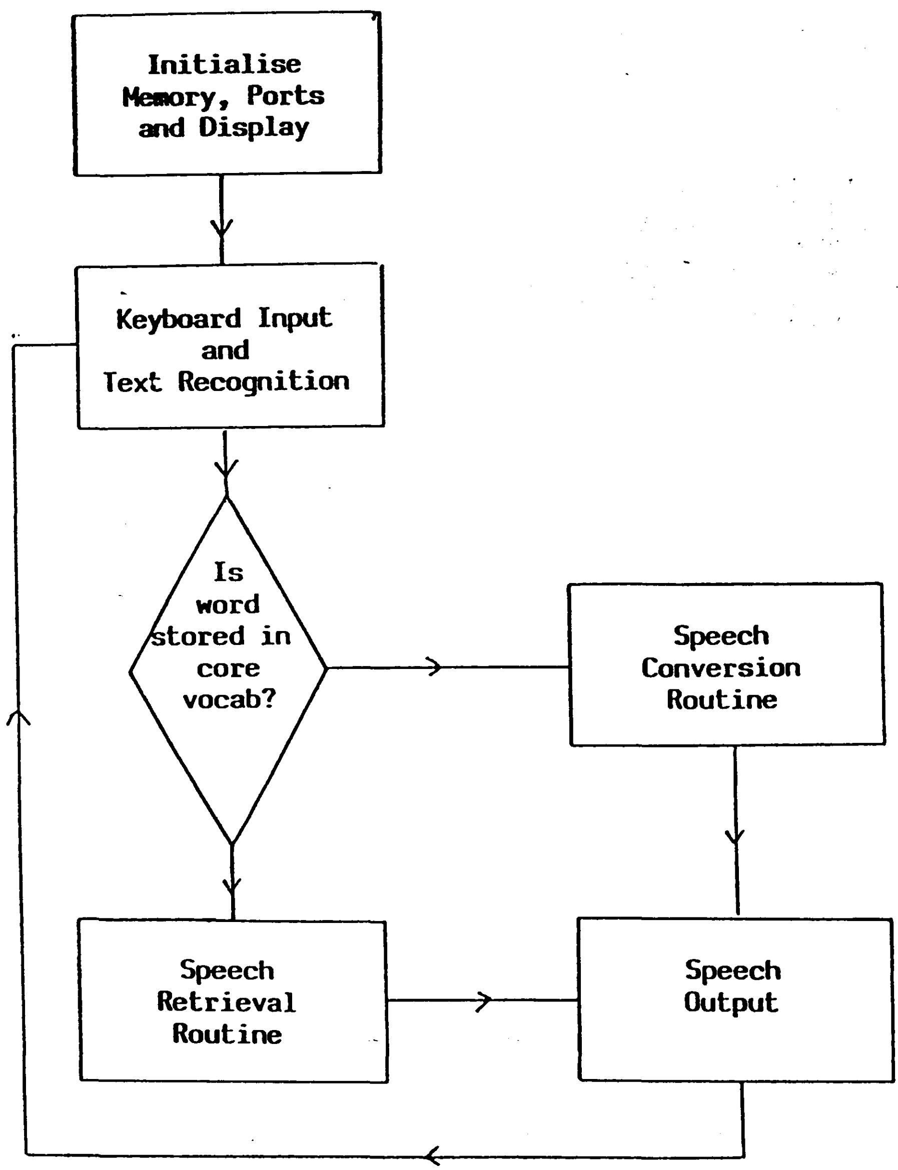

FIGURE 6.5 The Uvocom Software

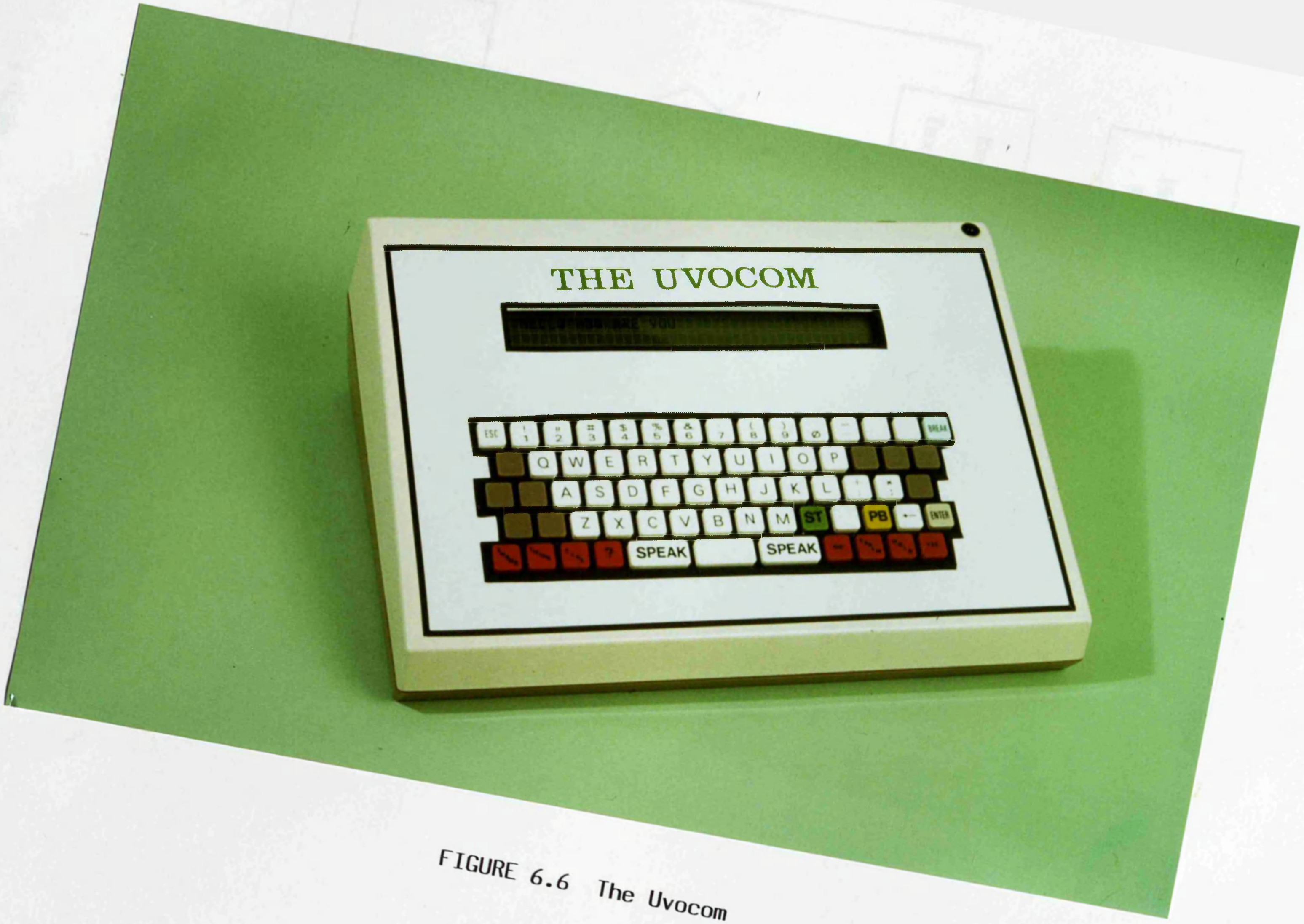

FIGURE 6.6 The Uvocom

written in 6502 assembly language to allow ease of modification for the transfer of the software from the BBC microcomputer to the specific hardware of the Uvocom.

There are three main software routines (Fig 6.5).

The main program loop is responsible for the initialisation of the memory, ports and display. This routine performs text recognition, display of text and determines whether the speech conversion or speech retrieval routine is called.

The speech conversion routine is called if the word entered on the keyboard is not stored in the core vocabulary. The text string is converted into a phoneme string and the commercial software then converts the phoneme code into LPC code to drive the 5220 synthesiser.

The speech retrieval routine is called if the entered word is stored in the core vocabulary. This routine sends data from EPROM to the speech processor.

A detailed description of the Uvocom software is given in Appendix 9.

6.4.4 Method of Use

The Uvocom is shown in Figure 6.6 and it can be used in two different modes.

a) **Interactive Mode**

In this mode of operation the SPEAK key is pressed after each word is typed in. This has the advantage of keeping the attention of the listener and gets round the problem of lack of intonation and also the slight difference in sound due to the two speech methods used. Communication breakdown is known to occur in other devices because of the long message preparation time.

The conversation prompting phrases are available using the red buttons so conversation prompting methods can be used supplemented by the typing of words or phrases. It is important to point out that the device should be used to convey meaning rather than to produce a grammatically correct dialogue.

b) **The Recall Mode**

Words or phrases can be stored by pressing the green ST button. A prompt is then displayed to input a code identifying letter, which can be any single character A to Z, followed by the word or phrase to be stored. The limit to the number of characters which can be stored in any phrase is 50 and the maximum number of phrases is 26 (one for each letter of the alphabet).

Stored phrases are recalled by pressing the yellow PB button followed by the code letter for the appropriate phrase.

If a phrase is given a code letter identical to the previously stored phrase the new phrase will replace the previous phrase.

6.5 Evaluation

Initial trials with the Uvocom device have been conducted with normals and with a limited number of patients. The feedback from these trials has been very encouraging and the results are summarised below.

6.5.1 Physical Format

Most users of the Uvocom found the device easy to handle and stable in use. One patient commented on the plastic front panel cover which was easily wiped clean when drooled over.

Some users found the quality of the QWERTY keyboard to be less than ideal. Sometimes a string of letters appeared which was due to 'key-bounce'. The Liquid Crystal Display was acceptable to most users but one user found it difficult to read in some lighting conditions.

6.5.2 Use of the Device

Users did not have any difficulty operating the device. The information sheets provided were adequate for this purpose. Recall mode was often used to prepare phrases in advance. Users had no difficulty in storing and recalling phrases. Some users commented on the need for battery-backed up RAM as at the moment stored phrases are lost when the device is switched off.

The interactive mode was used more often. Two methods were used. In the word/speak mode each word is spoken as it is entered and in the phrase/speak mode the phrase is spoken after it is entered. The word/speak mode was intended to keep listeners interested and this was indeed the case. One motor neurone disease patient wrote:

"In the word/speak mode, listeners hung on my every word, gasping for the total chat and in the phrase/speak mode, they were less impatient but were more often left wondering just what was being said".

The red Conversation Prompting buttons were not as successful as in the Pocket Speech Aid. This was partly due to users not being taught the techniques of conversation prompting and the quality of the recordings were poorer than the ADPCM Pocket Speech Aid.

The words stored in the core vocabulary were of better quality than the constructed words. Users were impressed by the size of the core vocabulary but many felt it necessary to experiment with different spellings for the constructed words.

When the Uvocom was presented to patients it was stressed that the device should be used to convey meaning rather than real dialogue and many individuals found it useful to use interjections such as 'Good', 'Sorry', 'Please' and 'Oh' - one patient used these interpolations to enable her to contribute to a telephone conversation.

6.5.3 The Communication Rate

One of the main drawbacks to using an unlimited vocabulary device is the slow communication rate. There are many techniques to increase the communication rate. One of the simplest methods is to stress to patients the need to convey meaning rather than proper dialogue. The techniques of conversation prompting can be applied to increase the communication rate.

The QWERTY keyboard is not the optimum keyboard for disabled users and keyboard optimisation techniques have been developed to increase communication rate. Keyboards have been developed that can reduce average between-key movement distances by 50% with predicted communication rate improvements of up to 35% (46). Damper and his colleagues at Southampton University have successfully used a similar technique in developing a keyboard for a hand-held text-to-speech device (47).

Many workers have developed techniques for word prediction based on previously entered words. Newell and his colleagues in Dundee have developed Predictive Adaptive Learning techniques (48) and a similar system has been developed in Glasgow for single switch users (49). These systems reduce the number of key-presses required to select a word or phrase. However, there are possible drawbacks to using a system such as this as the mental load placed on patients in increased as they are required to decide after each prediction whether the machine proposed character is what was intended or not. Thus while prediction may bring about a saving in keystrokes per word, it is responsible for increasing mental load and therefore may not increase the communication rate (50, 51).

6.5.4 Further Development

The initial trials with this device have indicated that further development is worthwhile and if this is undertaken then the device could become a useful communication aid. There are many improvements that can be made. Battery back up RAM can be added to enable long term phrase storage. A better quality keyboard could be used.

The overall speech quality could be improved by making refinements to the test to speech algorithm and by re-recording some of the words and phrases in the core vocabulary.

Further work can be done on improving the communication rate. Experience from clinical trials has shown that this can be achieved through teaching and use of conversation prompting techniques. The keyboard layout could be changed to suit disabled users but further work is necessary before incorporating predictive learning techniques. It is possible that further software refinements will be added to the Uvocom.

CHAPTER 7 - CONCLUSIONS

7.1 Original Aim

The original aim of this work was to develop a range of synthetic speech units which would be of practical benefit to neurological patients with speech impairment. These aids were to be modular in concept and clinical experience was to be used in the first instance to modify and adapt the basic design. The final objective was to develop aids with speech output which would be available at reasonable cost and to consider the practical aspects of providing such aids.

7.2 The Technical Features of the Devices

The devices were designed to cover a wide range of physical skills. Good finger movement is required for the Uvocom, limited finger control is sufficient for the Macleod Unit and only single finger operation is necessary for the Pocket Speech Aid.

The speech aids were designed as portable aids to enable individuals to use them in many different situations. Patients were not restricted to using the devices at home or in a hospital ward which is the case with many commercial systems that require a mains power source. The intelligibility and speech quality trials guided the type of synthetic speech which was to be used in each device. The method of synthesis with the highest speech quality was used to suit each application. Clinical trials have shown the importance of the portability and speech quality features. Synthetic speech is a poor substitute for normal speech and cannot hope to benefit a patient in every situation. Nevertheless, this

MEDICAL DIAGNOSIS	NO. OF CASE STUDIES	INCREASE IN COMMUNICATION	NO INCREASE IN COMMUNICATION
Cerebrovascular Accident	16	11	5
Motor Neurone Disease	14	7	7
Pseudobulbar Palsy	4	2	2
Multiple Sclerosis	2	1	1
Functional Aphonia	1	1	0
Parkinson's Disease	1	0	1
Unknown Etiology	1	1	0

TABLE 7.1 Case Study Summary

work has shown that synthetic speech can improve a speech impaired individual's ability to communicate in certain situations. The situations are different for each individual and it is therfore of immense importance that the device be portable and battery powered. The quality of the synthetic speech has proved to be of an acceptable standard and has rarely been the cause of rejection of a device. This was not the case for the commercial speech aids.

7.3 The Case Studies

Thirty-nine case studies were conducted and 23 patients increased their communication ability when using one of the speech aids. It is informative to examine the medical diagnosis of the patients studied in the project and this is illustrated in Table 7.1. Chapter 1 examined the possible patient groups that might benefit from a synthetic speech aid. Patients with motor neurone disease have been studied but patients suffering from other progressive disorders such as Fredreich`s Ataxia, Myasthenia Gravis and Huntington's Chorea have not been referred. Only two patients with Multiple Sclerosis and one patient with Parkinson's Disease have been assessed. Many patients who have suffered a Cerebrovascular Accident have been assessed but patients with other non-degenerative diseases such as the head injured have not been referred. In some cases patients with these diseases are not referred to Speech Therapy Departments. This may be due to clinical staff not being aware that alternative and augmentative forms of communication can be of benefit to them. This may be due to the small numbers of published case studies in this field. The results from this work suggest that other patient groups can benefit from synthetic speech and it is intended that further assessment be conducted at the Scottish Centre of Technology for the Communication Impaired (SCTCI).

7.4 **Presentation of a Speech Aid**

Presentation of a speech aid is extremely important. Patients must be taught how to use the device. The therapist can use many different techniques to help a patient make optimum use of a device. In this work the speech therapist made extensive use of instruction booklets, audio and video tapes to demonstrate the use of a device. Tapes can be used to illustrate to a patient that the device is most effective when supplementing other existing methods of communication. In some instances such as in the use of a Conversation Prompter, it is desirable that instruction is given to communication partners as well as to the user of the device. This could have far reaching implications in terms of speech therapists' time and it may be beneficial for more therapists to specialise in the teaching of alternative and augmentative communication strategies. Learning time is an important consideration and this should be kept to a minimum particularly for patients with progressive degenerative diseases. This is illustrated in case study 26. Before this patient had learnt to use the device his physical skills had deteriorated to such an extent that a simpler device was required. These patients are more likely to use a device that they can work easily. Therapy sessions on the use of a device can help in reducing the learning time.

A device should not be issued to a patient without proper instruction on its use and this work has highlighted the need for further resources to be directed towards teaching methods and device presentation.

7.5 Technology Transfer

Several Communication Aids Centres were visited during the course of this work. These centres provide an assessment facility for speech impaired patients and recommendations are made for suitable communication aids. These visits showed that the price and availability of devices was a recurrent problem. There are some speech aids available which are too expensive and can only be made available to a very small number of patients. Some of the clinical disorders discussed in Chapter 1 are degenerative and these patients may require several different devices as the disease progresses. For these patients it would be appropriate for devices to be lent out from a central lending bank such as the SCTCI.

The devices developed in this project are low cost devices which has been the key feature in the technology transfer process. Where possible, commercial equipment should be used or adapted to meet particular applications. This has been achieved by many workers but unfortunately this is more difficult when portability is an important feature. The parts cost of the Pocket Speech Aid was kept to a minimum (£90). This helped to attract commercial interest. An agreement for the Pocket Speech Aid was signed with a commercial company on 14 October 1986 but this venture proved to be unsuccessful. The company showed early initial enthusiasm but for the wrong reasons. They wanted to be involved in this emotive area so that the product could be included in a portfolio which was to be presented to merchant bankers. Progress with this company was very slow and the licensing agreement was discontinued by Greater Glasgow Health Board.

A new licensing agreement was signed with another company on 19 February 1988 and the technology transfer should be achieved more easily because this company has more experience in the medical field.

The Pocket Speech Aid will retail at approximately £300 and should be inexpensive enough to enable Communication Aids Centres throughout the UK to purchase a reasonable number of units.

A recording station based on a plug-in card for the IBM personal computer has been developed in the Department of Clinical Physics and Bio-Engineering. This card will be commercially available from the same company and this should enable Communication Aid Centres throughout the United Kingdom and possibly comparable centres overseas to record and program phrases for the Pocket Speech Aids in use in their own area.

One important feature in this approach is that referral would always be through speech therapists who would have responsibility and control over the selection and recording of phrases.

The Uvocom device also has potential for commercial development. Based on encouraging clinical feedback this device will undergo extended clinical trials before a decision on marketing is reached. The parts cost for the Uvocom is approximately £200 and it would be expected that a commercial device would cost about £600 to produce.

7.6 Future Work

There are two directions in which further work could progress. The techniques of prescription and assessment could be investigated and further development of the devices could be undertaken.

7.6.1 Assessment

In this work the assessment of the devices has been restricted to a set of case studies. A literature search, visits to Communication Aid Centres both in the UK and in the USA and attendance at international conferences has highlighted the need for a more scientific approach to assessment. At present, the success in matching the individual patient to a communication aid is subjectively based on the skill and experience of the speech therapist. A multi-disciplinary approach and more resources are needed to produce an objective method of prescription and assessment.

The process of assessing speech impaired persons for the prescription of a communication aid requires the determination of physical and cognitive abilities together with definition of needs and goals. Standard tests could be used to determine these features, for example, cognitive abilities can be assessed by means of parts of the Boston Diagnostic Aphasia Examination (BDAE), the Weschler Adult Intelligence Scale (WAIS Block Design) and the Ravens Matrices tests. A database could be developed which would include details of each communication aids features and modes of use. This database could be used to match a device to a patient's skills, abilities, needs and goals. Methods of cataloguing and analysis have already been suggested by Rosen and Goodenough

Trepagnier (52).

A more difficult aspect of assessment is the retrospective assessment of the improvement in the person's communication ability. Current approaches favour functional communication profiles which are recorded before an aid is prescribed and after a suitable period of familiarisation with the augmentative system (53, 54).

7.6.2 Further Development of the Devices

The Pocket Speech Aid and the Uvocom are two devices developed in this project that could generate further development work. A further extension of the Pocket Speech Aid concept could be the introduction of a RAM based system which could incorporate a hand held recording facility. The cost of RAM chips has fallen and the memory capacity increased. This would further simplify recording of phrases and would be an alternative to the ROM based Pocket Speech Aid. The parts cost for a RAM based system would be approximately £200. The Uvocom is an unlimited vocabulary device which uses LPC recording and phoneme concentration. There are other techniques that could provide another alternative to this approach.

Foulds and co-workers in the USA are developing a diphone system for a personal microcomputer (55). With CMOS Digital Signal Processing Chips becoming available it is now possible to develop a portable diphone synthesis machine. The diphone inventory for the DECTALK produces American sounding speech which may be less acceptable to users than the Uvocom speech. If a portable unlimited vocabulary device which uses diphones were to be developed a new diphone inventory may be required to provide patients with a more suitable accent and dialect.

7.7 Summary

Patients suffering from acquired speech loss and in particular those with progressive degenerative diseases are faced with many physical and emotional problems. Speech loss is one of the most distressing symptoms of their disease because they are faced with the inability to communicate with their family and friends - something they have been used to all their lives.

This work has investigated the use of synthetic speech as a method of improving the communication skills of these patients and it has been shown that many individuals can indeed increase their communication ability with the help of a speech aid. There are many problems still to be solved such as device prescription and patient assessment, but work is progressing in these difficult areas. Technology is also improving which will result in improved speech aids with more natural sounding speech and faster communication rates. This work has shown that technology can benefit this group of the disabled population and it is worthwhile for further effort and resources to be directed to improving the communication skills of patients with acquired speech loss.

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	87.5	100.0	100.0	100.0	100.0	100.0	97.0
Subject 2	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 3	100.0	100.0	100.0	87.5	100.0	100.0	97.9
Subject 4	100.0	100.0	100.0	87.5	100.0	100.0	97.9
Subject 5	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 6	100.0	100.0	100.0	87.5	100.0	100.0	97.9
Subject 7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 8	87.5	100.0	87.5	100.0	100.0	100.0	95.8
Subject 9	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 10	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 11	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 12	87.5	100.0	100.0	87.5	100.0	100.0	95.8
Subject 13	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 14	87.5	100.0	100.0	100.0	87.5	100.0	95.8
Subject 15	87.5	100.0	100.0	87.5	100.0	100.0	95.8
Subject 16	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 17	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 18	100.0	100.0	100.0	87.5	100.0	100.0	97.9
Subject 19	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 20	87.5	100.0	100.0	75.0	100.0	100.0	93.7
Subject 21	87.5	100.0	100.0	87.5	100.0	100.0	95.8
Subject 22	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 23	100.0	100.0	100.0	75.0	100.0	100.0	95.8
Subject 24	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 25	100.0	100.0	100.0	100.0	100.0	100.0	100.0

TABLE A1a.1 DRT Scores for Male Control

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 2	87.5	100.0	100.0	100.0	87.5	100.0	95.8
Subject 3	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 4	75.0	100.0	100.0	100.0	62.5	100.0	89.6
Subject 5	87.5	100.0	100.0	100.0	87.5	100.0	95.8
Subject 6	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 8	87.5	100.0	100.0	87.5	75.0	100.0	91.7
Subject 9	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 10	87.5	100.0	100.0	100.0	87.5	100.0	95.8
Subject 11	87.5	100.0	100.0	100.0	75.0	100.0	93.7
Subject 12	87.5	100.0	100.0	87.5	62.5	87.5	87.5
Subject 13	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 14	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 15	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 16	100.0	100.0	100.0	87.5	100.0	100.0	97.9
Subject 17	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 18	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 19	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 20	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 21	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Subject 22	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 23	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 24	100.0	100.0	100.0	100.0	87.5	100.0	97.9
Subject 25	100.0	100.0	100.0	100.0	87.5	100.0	97.9

TABLE A1a.2 DRT Scores for Female Control

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	100.0	100.0	100.00	75.0	87.5	100.0	93.7
Subject 2	100.0	100.0	87.5	25.0	87.5	100.0	83.3
Subject 3	100.0	100.0	87.5	62.5	50.0	100.0	83.3
Subject 4	100.0	100.0	87.5	37.5	62.5	100.0	81.2
Subject 5	100.0	100.0	75.0	62.5	87.5	100.0	87.5
Subject 6	100.0	100.0	100.0	25.0	75.0	87.5	81.2
Subject 7	100.0	100.0	87.5	50.0	75.0	100.0	85.4
Subject 8	87.5	100.0	87.5	50.0	75.0	100.0	83.3
Subject 9	100.0	100.0	75.0	62.5	75.0	100.0	85.4
Subject 10	100.0	100.0	75.0	50.0	87.5	100.0	85.4
Subject 11	87.5	100.0	100.0	25.0	75.0	87.5	79.2
Subject 12	100.0	100.0	87.5	50.0	62.5	87.5	81.2
Subject 13	100.0	100.0	100.0	75.0	100.0	87.5	93.7
Subject 14	87.5	100.0	75.0	75.0	75.0	100.0	85.4
Subject 15	100.0	100.0	87.5	62.5	87.5	100.0	89.6
Subject 16	87.5	100.0	75.0	62.5	100.0	87.5	85.4
Subject 17	87.5	100.0	87.5	62.5	75.0	100.0	85.4
Subject 18	100.0	100.0	100.0	62.5	62.5	100.0	87.5
Subject 19	100.0	100.0	62.5	50.0	87.5	100.0	83.3
Subject 20	100.0	100.0	100.0	37.5	87.5	100.0	87.5
Subject 21	100.0	100.0	100.0	62.5	87.5	100.0	91.7
Subject 22	100.0	100.0	100.0	62.5	87.5	87.5	89.6
Subject 23	100.0	100.0	87.5	75.0	87.5	87.5	89.6
Subject 24	75.0	100.0	100.0	37.5	100.0	100.0	85.4
Subject 25	100.0	100.0	87.5	50.0	62.5	100.0	83.3

TABLE A1a.3 DRT Scores for Male ADPCM

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	100.0	100.0	87.5	87.5	75.0	75.0	87.5
Subject 2	100.0	100.0	100.0	75.0	62.5	75.0	85.4
Subject 3	100.0	87.5	75.0	75.0	62.5	87.5	81.2
Subject 4	100.0	100.0	75.0	75.0	50.0	75.0	79.2
Subject 5	100.0	100.0	100.0	75.0	75.0	50.0	83.3
Subject 6	100.0	100.0	100.0	87.5	75.0	87.5	91.7
Subject 7	87.5	100.0	87.5	87.5	37.5	100.0	83.3
Subject 8	100.0	100.0	87.5	75.0	25.0	75.0	77.1
Subject 9	87.5	100.0	62.5	87.5	0.0	62.5	66.7
Subject 10	75.0	100.0	87.5	75.0	75.0	75.0	81.2
Subject 11	100.0	100.0	75.0	75.0	50.0	62.5	77.1
Subject 12	87.5	100.0	75.0	75.0	37.5	62.5	72.9
Subject 13	75.0	100.0	87.5	87.5	37.5	75.0	77.1
Subject 14	87.5	100.0	87.5	87.5	50.0	75.0	81.2
Subject 15	100.0	100.0	87.5	87.5	62.5	75.0	85.4
Subject 16	75.0	87.5	62.5	87.5	75.0	75.0	77.1
Subject 17	100.0	100.0	87.5	75.0	37.5	75.0	79.2
Subject 18	75.0	100.0	75.0	75.0	62.5	37.5	70.8
Subject 19	50.0	100.0	75.0	50.0	50.0	87.5	68.7
Subject 20	87.5	100.0	100.0	75.0	0.0	37.5	66.7
Subject 21	87.5	100.0	75.0	75.0	37.5	62.5	72.9
Subject 22	75.0	100.0	87.5	75.0	0.0	87.5	70.8
Subject 23	100.0	100.0	100.0	75.0	50.0	87.5	85.4
Subject 24	87.5	100.0	100.0	75.0	50.0	87.5	83.3
Subject 25	100.0	100.0	87.5	75.0	37.5	50.0	75.0

TABLE A1a.4 DRT for Female ADPCM

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	100.0	87.5	62.5	75.0	87.5	62.5	79.2
Subject 2	87.5	87.5	62.5	100.0	87.5	75.0	83.3
Subject 3	100.0	87.5	85.0	75.0	75.0	87.5	85.4
Subject 4	87.5	100.0	100.0	75.0	50.0	75.0	81.2
Subject 5	100.0	87.5	75.0	87.5	87.5	75.0	85.4
Subject 6	100.0	75.0	62.5	87.5	100.0	62.5	81.2
Subject 7	100.0	100.0	50.0	100.0	62.5	87.5	83.3
Subject 8	100.0	87.5	75.0	75.0	50.0	75.0	77.1
Subject 9	75.0	100.0	75.0	75.0	62.5	75.0	77.1
Subject 10	87.5	100.0	62.5	87.5	75.0	75.0	81.2
Subject 11	87.5	75.0	62.5	75.0	62.5	62.5	70.8
Subject 12	87.5	87.4	50.0	75.0	37.5	87.5	70.8
Subject 13	100.0	100.0	87.5	87.5	62.5	100.0	89.6
Subject 14	100.0	87.5	75.0	87.5	75.0	50.0	79.2
Subject 15	100.0	75.0	62.5	62.5	87.5	100.0	81.2
Subject 16	87.5	75.0	75.0	87.5	75.0	87.5	81.2
Subject 17	87.5	100.0	62.5	87.5	75.0	75.0	81.2
Subject 18	100.0	100.0	37.5	62.5	87.5	75.0	77.1
Subject 19	100.0	100.0	62.5	87.5	87.5	62.5	83.3
Subject 20	100.0	75.0	75.0	62.5	75.0	75.0	77.1
Subject 21	100.0	100.0	87.5	62.5	50.0	87.5	81.2
Subject 22	100.0	100.0	50.0	75.0	62.5	62.5	75.0
Subject 23	100.0	100.0	62.5	75.0	75.0	75.0	81.2
Subject 24	87.5	87.5	75.0	75.0	75.0	75.0	79.2
Subject 25	87.5	100.0	62.5	75.0	87.5	87.5	83.3

TABLE A1a.5 DRT for Male LPC

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	100.0	100.0	75.0	62.5	87.5	62.5	81.2
Subject 2	100.0	100.0	62.5	50.0	62.5	87.5	77.1
Subject 3	100.0	100.0	62.5	62.5	75.0	62.5	77.1
Subject 4	87.5	87.5	62.5	87.5	37.5	75.0	72.9
Subject 5	87.5	100.0	37.5	75.0	62.5	50.0	68.7
Subject 6	87.5	100.0	75.0	62.5	62.5	87.5	79.2
Subject 7	100.0	100.0	75.0	62.5	37.5	75.0	75.0
Subject 8	75.0	100.0	62.5	75.0	63.5	75.0	75.0
Subject 9	100.0	100.0	87.5	75.0	75.0	100.0	89.6
Subject 10	100.0	100.0	50.0	75.0	50.0	75.0	75.0
Subject 11	87.5	100.0	62.5	87.5	75.5	87.5	83.3
Subject 12	100.0	100.0	75.0	75.0	37.5	87.5	79.2
Subject 13	100.0	100.0	75.0	87.5	62.5	62.5	81.2
Subject 14	100.0	100.0	37.5	50.0	50.0	75.0	68.7
Subject 15	100.0	100.0	50.0	62.5	62.5	75.0	75.0
Subject 16	100.0	100.0	62.5	37.5	62.5	62.5	70.8
Subject 17	87.5	100.0	50.0	75.0	37.5	50.0	66.7
Subject 18	100.0	100.0	75.0	87.5	12.5	100.0	79.2
Subject 19	87.5	100.0	87.5	50.0	50.0	75.0	75.0
Subject 20	100.0	87.5	75.0	50.0	25.0	75.0	68.7
Subject 21	87.5	62.5	37.5	75.0	62.5	87.5	68.7
Subject 22	87.5	100.0	62.5	50.0	50.0	75.0	70.8
Subject 23	100.0	87.5	75.0	50.0	37.5	37.5	64.6
Subject 24	100.0	100.0	75.0	62.5	75.0	75.0	81.2
Subject 25	100.0	100.0	75.0	62.5	75.0	75.0	81.2

TABLE A1a.6 DRT Scores for Female LPC

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	62.5	50.0	50.0	62.5	12.5	50.0	47.9
Subject 2	62.5	87.5	62.5	62.5	75.0	37.5	64.6
Subject 3	75.0	62.5	75.0	62.5	50.0	62.5	64.6
Subject 4	37.5	37.5	25.0	50.0	87.5	25.0	43.7
Subject 5	50.0	62.5	100.0	75.0	37.5	50.0	62.5
Subject 6	75.0	62.5	62.5	50.0	50.0	62.5	60.4
Subject 7	75.0	50.0	75.0	62.5	25.0	25.0	52.1
Subject 8	12.5	62.5	50.0	62.5	37.5	25.0	41.7
Subject 9	62.5	25.0	37.5	62.5	62.5	62.5	52.1
Subject 10	12.5	75.0	50.0	75.0	25.0	50.0	47.9
Subject 11	75.0	62.5	87.5	75.0	12.5	37.5	58.3
Subject 12	12.5	25.0	62.5	62.5	62.5	0.0	37.5
Subject 13	37.5	75.0	62.5	25.0	25.0	62/5	47.9
Subject 14	75.0	50.0	75.0	75.0	12.5	25.0	52.1
Subject 15	50.0	62.5	87.5	62.5	50.0	62.5	62.5
Subject 16	25.0	62.5	87.5	50.0	25.0	37.5	47.9
Subject 17	37.5	75.0	87.5	62.5	62.5	37.5	60.4
Subject 18	62.5	62.5	100.0	37.5	37.5	62.5	60.4
Subject 19	50.0	62.5	50.0	12.5	50.0	37.5	43.7
Subject 20	87.5	37.5	75.0	62.5	50.0	0.0	52.1
Subject 21	75.0	25.0	62.5	62.5	25.0	62.5	52.1
Subject 22	75.0	100.0	100.0	50.0	50.0	50.0	70.8
Subject 23	50.0	62.5	62.5	50.0	62.5	25.0	52.1
Subject 24	37.5	37.5	37.5	50.0	37.5	37.5	39.6
Subject 25	75.0	50.0	62.5	87.5	50.0	37.5	60.4

TABLE A1a.7 DRT Scores for Votrax

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	DRT score
Subject 1	50.0	75.0	100.0	0.0	87.5	100.0	68.7
Subject 2	37.5	50.0	87.5	12.5	87.5	62.5	56.2
Subject 3	25.0	100.0	100.0	-37.5	100.0	87.5	62.5
Subject 4	37.5	75.0	62.5	-25.0	87.5	62.5	50.0
Subject 5	50.0	62.5	87.5	0.0	62.5	87.5	58.3
Subject 6	50.0	75.0	87.5	-12.5	50.0	87.5	56.2
Subject 7	50.0	62.5	75.0	-12.5	75.0	100.0	58.3
Subject 8	37.5	87.5	62.5	12.5	87.5	87.5	62.5
Subject 9	62.5	50.0	25.0	25.0	87.5	100.0	58.3
Subject 10	25.0	37.5	37.5	-37.5	87.5	87.5	39.6
Subject 11	25.0	62.5	62.5	-12.5	100.0	75.0	52.1
Subject 12	37.5	87.5	87.5	0.0	50.0	50.0	52.1
Subject 13	87.5	75.0	100.0	0.0	62.5	75.0	66.7
Subject 14	37.5	50.0	50.0	-12.5	75.0	87.5	47.9
Subject 15	50.0	75.0	75.0	12.5	87.5	87.5	64.6
Subject 16	50.0	75.0	87.5	0.0	50.0	75.0	56.2
Subject 17	87.5	62.5	87.5	0.0	75.0	87.5	66.7
Subject 18	62.5	12.5	37.5	25.0	62.5	75.0	45.8
Subject 19	0.0	62.5	75.0	-12.5	87.5	100.0	52.1
Subject 20	50.0	50.0	87.5	-25.0	75.0	87.5	54.2
Subject 21	25.0	50.0	75.0	25.0	75.0	87.5	56.2
Subject 22	62.5	62.5	62.5	12.5	12.5	50.0	43.7
Subject 23	62.5	50.0	50.0	12.5	62.5	62.5	50.0
Subject 24	37.5	87.5	50.0	12.5	87.5	87.5	60.4
Subject 25	62.5	100.0	87.5	37.5	75.0	87.5	75.0

TABLE A1a.8 DRT Scores for SAM

	Voicing	Nasality	Sustension	Sibilation	Graveness	Compactness	Overall
Male Control	100	100	100	100	100	100	100
Female Control	100	100	100	100	87.5	100	100
Male ADPCM	100	100	87.5	62.5	87.5	100	87.5
Female ADPCM	87.5	100	87.5	75	50	75	81.25
Male LPC	100	87.5	62.5	75	75	75	75
Female LPC	100	100	62.5	62.5	62.5	75	75
Votrax	62.5	62.5	62.5	62.5	50	37.5	56.25
SAM	50	62.5	75	0	75	87.5	62.5

TABLE A1a.9 Median DRT Scores

APPENDIX 1a - THE INTELLIGIBILITY TRIAL

Intelligibility

A Diagnostic Rhyme Test (DRT) score was found for each tape and individual scores for each distinctive feature. The scores are adjusted for guessing using the following correction.

$$\text{DRT Score} = \frac{100\ (\text{correct} - \text{incorrect})}{\text{Total}}$$

As the data are not normally distributed about a mean value, two sample, non parametric comparisons of medians were made using Wilcoxin-Mann-Whitney tests. The objective is to test whether the two groups of data are centred differently. The DRT scores are shown in Tables A1a.1 - A1a.8. Median intelligibility scores are extracted from these tables and are shown in Table A1a.9.

For each group comparison the null hypothesis that the two sets of data are identical is tested. The groups of data are ranked in terms of intelligibility scores and the Wilcoxin Statistic, W, is defined as the sum of the ranks in one of the groups. The lower value of W, the stronger the evidence for rejecting the null hypothesis. The sample size is larger than those covered by the Wilcoxin table and W can be considered to have a normal distribution (54).

The distribution of the test statistic W has a mean value:

$$E\ (W) = m\ (m + n + 1)$$

where m = number of results in the first group, and n = number of results in the second group, and a variance of:

$$\text{Var}\ (W) = \frac{1}{12}\ mn\ (m + n + 1)$$

Probability values can be obtained by referring the critical ratio:

$\frac{W - E(W)}{\sigma_W}$ to normal probability tables.

where σ_W = standard deviation of W.

The following results are obtained at the 1% level of significance.

1) The control data is more intelligible than the synthetic methods.
2) Male ADPCM speech is more intelligible than Male LPC speech.
3) Male and female LPC speech are more intelligible than Votrax and SAM.
4) Male ADPCM speech is more intelligible than female ADPCM speech.

The following results are obtained at the 5% level of significance.

1) SAM is more intelligible than Votrax.
2) There is nc significant difference between female ADPCM and female LPC speech.
3) There is no significant difference between male LPC and female LPC speech.

TESTED	A v B		A v C		A v D		B v C		B v D		C v D	
PREFERRED	A	B	A	C	A	D	B	C	B	D	C	D
Observer 1	6	2	0	8	0	8	0	8	0	8	8	0
2	5	3	0	8	3	5	0	8	1	7	8	0
3	7	1	0	8	6	2	0	8	1	7	8	0
4	6	2	0	8	1	7	1	7	1	7	6	2
5	6	2	2	6	6	2	1	7	1	7	6	2
6	7	1	1	7	6	2	1	7	1	7	8	0
7	4	4	0	8	0	8	1	7	1	7	7	1
8	6	2	3	5	3	5	1	7	2	6	5	3
9	5	3	0	8	3	5	1	7	1	7	8	0
10	3	5	1	7	1	7	0	8	1	7	7	1
11	2	6	0	8	0	8	0	8	3	5	7	1
12	2	6	0	8	0	8	0	8	0	8	8	0
13	6	2	3	5	3	5	1	7	2	6	6	2
14	2	6	1	7	0	8	1	7	0	8	5	3
15	1	7	0	8	1	7	0	8	1	7	8	0
16*	7	1	1	7	6	2	7	1	2	6	7	1
17	6	2	2	6	4	4	0	8	0	8	6	2
18	5	3	0	8	1	7	0	8	0	8	4	4
19	5	3	0	8	1	7	0	8	0	8	8	0
20	5	3	1	7	4	4	4	4	3	5	3	5
TOTAL	96	64	15	145	49	111	19	141	21	139	133	27

(*) Listener 16's data were omitted from the second analysis - she was the only respondent to prefer method B to method C consistently.

A: Votrax, B: Text To Speech (BBC Micro), C: LPC words, D: Mixed LPC - Constructed words.

TABLE A1b.1 The preferences indicated by each listener (paired comparisons for each of 8 phrases).

Speech Quality Trial

The raw data from the record forms have been collated in Table A1b.1 and the last row of this table indicates that:

- method C was preferred to the other methods in almost 90% of the relevant comparisons;
- method D was preferred to B in almost 90% of cases and preferred to A in 70% of cases;
- method A was preferred to method B in 60% of the trials.

The standard statistical procedures that have been developed for paired-comparisons trials are based on the following model:

Let p (i,j) denote the probability that method i is preferred to method j in a paired comparisons trial. (p(i,j) + p(j,i)=1)

It is assumed that:

$$p(i,j) = \frac{p(i)}{p(i) + p(j)} \qquad \text{———— A1.1}$$

where p(i) (i=A, B, C, D) lies in the range (0, 1) and is interpreted as a relative preference parameter.

$$(p(A) + p(B) + p(C) + p(D) = 1)$$

The model contains only four parameters, the p(i) parameters.

To rank the methods it is necessary only to derive point estimates of the p(i) and then use these estimates to form a ranking for the parameters, for if p(i) is greater than p(j) then it is concluded that method i is preferred to method j.

Using the data given in table A1b.1 the following parameter estimates were determined:

PREFERRED METHOD	METHOD COMPARED				
	A	B	C	D	TOTAL
A	- (-)	96 (96)	15 (16)	49 (46)	160 (160)
B	64 (62)	- (-)	19 (10)	21 (32)	104 (104)
C	145 (144)	141 (150)	- (-)	133 (125)	419 (419)
D	111 (114)	139 (128)	27 (35)	- (-)	277 (277)

A: Votrax, B: Test To speech (BBC Micro), C: LPC words, D: Mixed LPC - Constructed words.

TABLE A1b.2 Total number of preferences in each paired comparison: Observed value and (Expected value under model).

p(C) = 0.6853

p(D) = 0.1907

p(A) = 0.0762

p(B) = 0.0477

It remained to determine whether the model (A1.1) actually fits the observed data sufficiently well to make these estimates a meaningful summary of the data. In this trial, each pair of methods was compared in a total of 160 tests, and so, under the model assumptions, the expected number of times that i was preferred to j in a paired comparison is:

$$160 \times \frac{p(i)}{p(i) + p(j)} \qquad \text{———} \qquad A1.2$$

These expected values were estimated by substituting in (A1.2) the parameter estimates given above. The estimated expected values and the corresponding observations are shown in table A1b.2.

A Chi-Square Goodness-of-Fit test was carried out using these values (52). Evidence was found to reject the Null Hypothesis that the model fitted the data well. Further examination of Table A1b.2 showed that the observed and expected values in most cells were in good agreement. The greatest discrepancy occured because B was, in practice, preferred much more often to C than expected. Tracing this back to Table A1b.1, this result is mainly due to the behaviour of listener 16, who recorded a preference for B over C in 7 out of 8 trials, in contrast to the other 19 listeners who recorded only 12 preferences for B in a total of 152 comparisons with C. In all other ways listener 16's responses seemed similar to those of the other respondents. It is reasonable to assume that, whatever the reason for this different pattern of preference between B and C, the same model is unlikely to

PREFERRED METHOD	METHOD COMPARED				
	A	B	C	D	TOTAL
A	- (-)	89 (94)	14 (13)	43 (40)	146 (147)
B	63 (58)	- (-)	12 (8)	19 (27)	94 (93)
C	138 (139)	140 (144)	- (-)	126 (121)	404 (404)
D	109 (112)	133 (125)	26 (31)	- (-)	268 (268)

A: Votrax, B: Text To Speech (BBC Micro), C: LPC words, D: Mixed LPC - Constructed words

TABLE A1b.3 Total number of preferences in each paired comparison (excluding listener 16's data): Observed value and (Expected value under model).

account for the tastes of listener 16 and the others. Consequently, this person's data were omitted and the remainder of the dataset re-analysed.

New estimates of the parameters were obtained:

p(C) = 0.7105

p(D) = 0.1839

p(A) = 0.0650

p(B) = 0.0405

Estimated expected values were again computed; these are given in table A1b.3. As expected, the discrepancy between the expected and observed values for comparisons between B and C had been reduced. When the Chi-Square Test was again carried out, there was now no evidence to reject the null hypothesis that the model fits the data.

The following 95% Confidence Intervals for p(A), p(B), p(C) and p(D) were then computed in the way recommended by Bradley (53):

p(C) : 0.6710 - 0.7500

p(D) : 0.1427 - 0.2251

p(A) : 0.0475 - 0.0825

p(B) : 0.0169 - 0.0641

The confidence interval for p(C) does not overlap that for p(D) therefore method C is indeed preferred to method D. Similarly, method D is preferred to methods A and B. On the other hand, although p(A) > p(B), the confidence intervals for these parameters overlap, so it is not possible to conclude with any certainty that A is generally preferred to B.

To try to resolve the uncertainty between A and B the data involving the other two methods are ignored and the individual preferences between A and B are analysed. The data in the first two columns of Table A1b.1 show that the majority of listeners (13 out of 19) preferred A to B for most of the 8 phrases with which they were presented. However, 5 out of 19 preferred B to A more often than A to B, and there was one tie. If it is assumed that an individual chosen at random from the population will prefer A to B for most of the 8 phrases with probability p then the Sign Test can be used to examine the two hypotheses:

H0 : $p=0.5$ (i.e A and B are each preferred by half the population)

H1 : $p>0.5$ (i.e. A is preferred by more than half the population)

When this test is applied to the observed data, it is not possible, at the 5% level of significance, to reject H0 in favour of H1, i.e it is again not possible to separate methods A and B as regards to general preference.

The final ranking of the four methods of speech production is:-

C > D > (A,B)

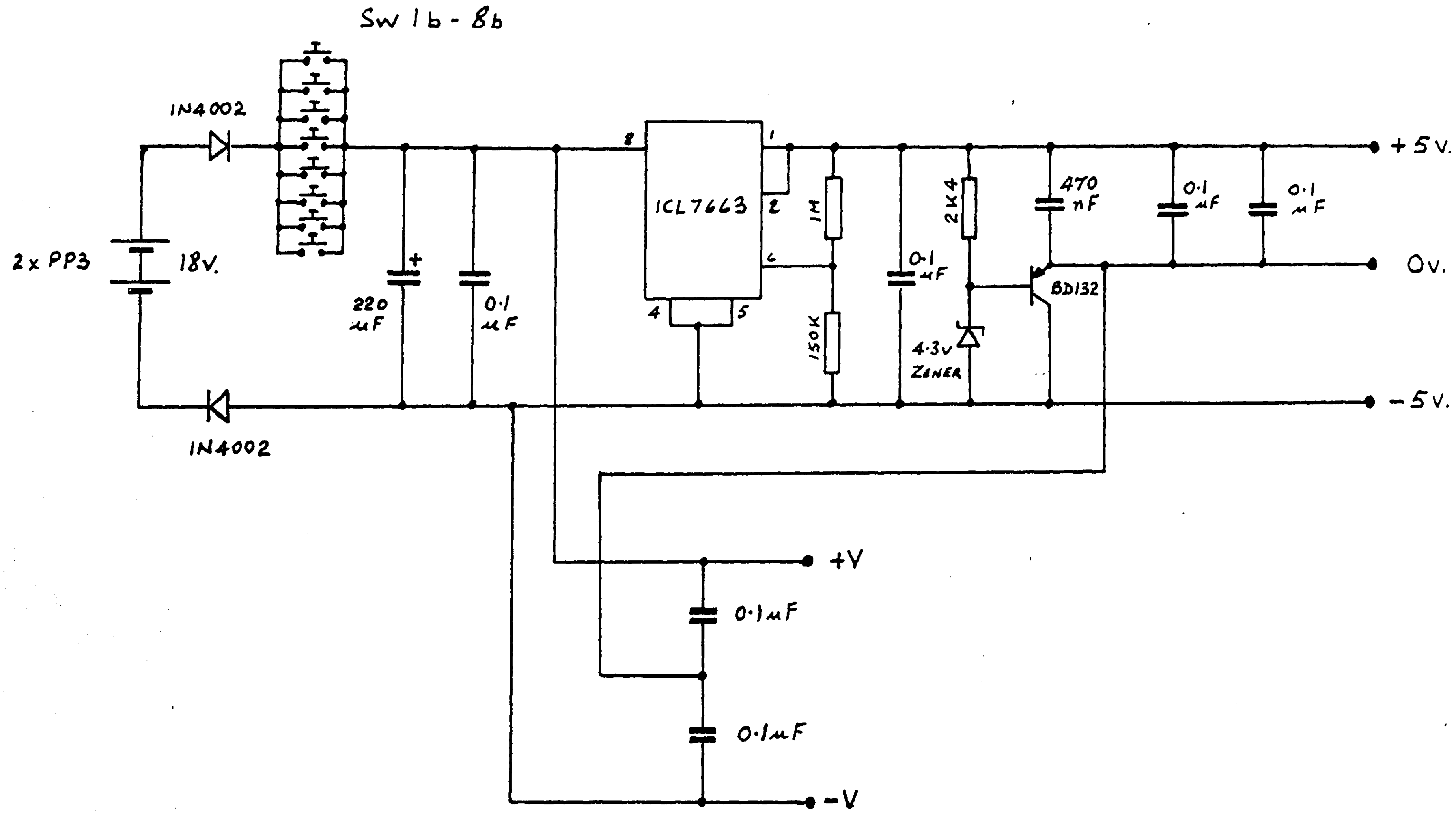

FIGURE A2.1 The Pocket Speech Aid Mk I: Power Supply

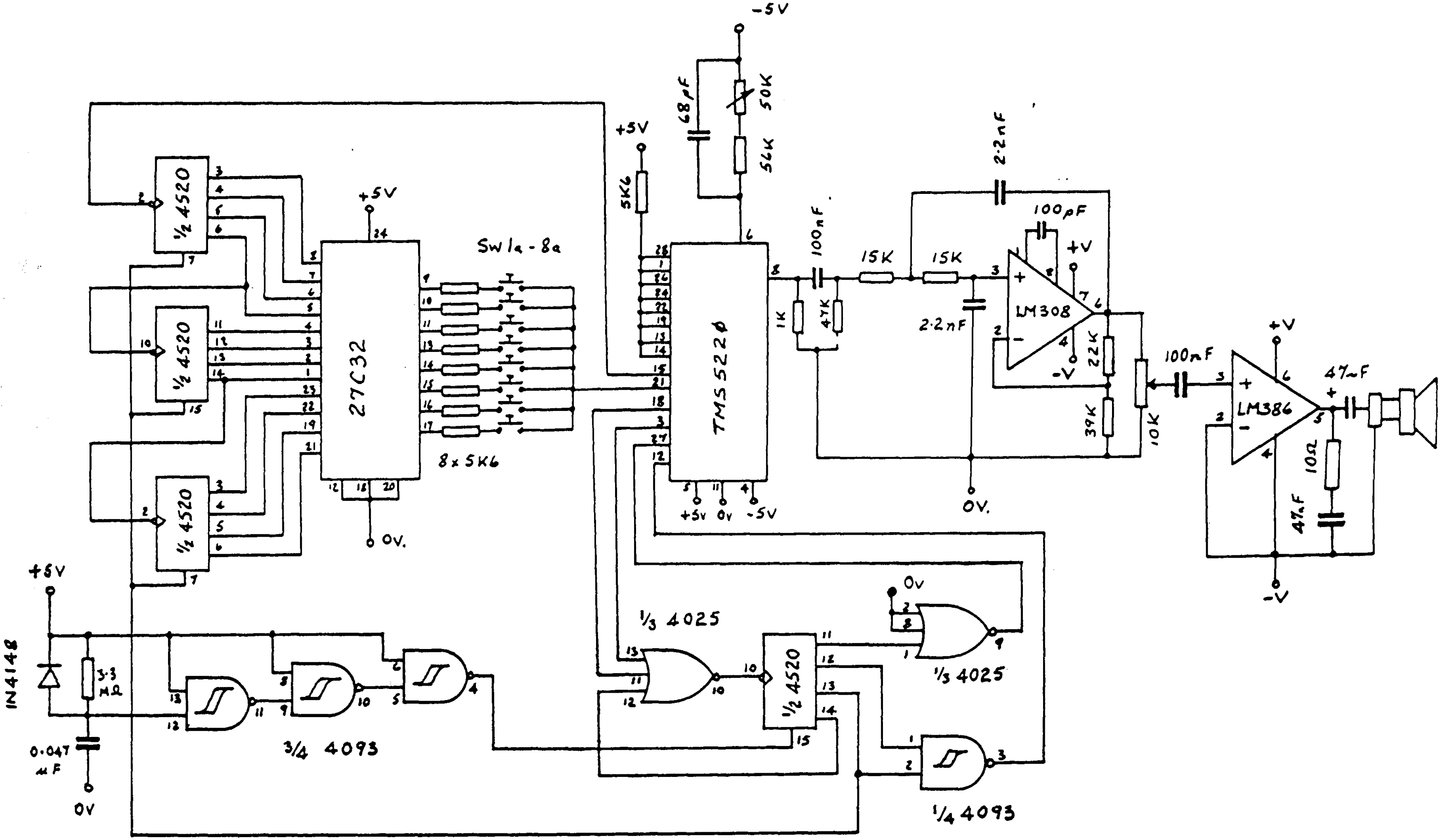

FIGURE A2.2 The Pocket Speech Aid Mk I: Circuit Diagram

APPENDIX 2 - THE POCKET SPEECH AID Mk I

Circuit Description

The power supply circuit of the device is shown in Figure A2.1 and the main circuit is shown in Figure A2.2. Power is applied to the circuit by the pressing of any one of the switches SW1 - 8. This will supply approximately 16V directly to the active filter and audio amplifiers. An ICL7663 voltage regulator and a zener diode, transistor configuration is used to produce a +5V supply. This supply is used by the CMOS logic, EPROM and the TMS 5220 speech synthesiser. In order to meet the TMS 5220 power-up clear requirements of +5V in less than 2 msec, decoupling capacitance on the logic supplies are kept to a minimum.

During power up an RC network and 3 gates on the 4093 cause the circuit to idle for a settling time of approximately 150 msec. The synthesiser chip is then issued, over its data bus, 3 RESET (FF) and a SPEAK (FB) command. These commands and the required hand-shake protocol are produced by the CMOS logic.

After 150 msec. the forced reset is then removed from the 4520 counter. The ROMCLK signal from the synthesiser chip is gated in a 4025 NOR gate with the synthesiser READY signal, WS causes the contents of data lines D0-D7 to be written into the synthesiser chip. Since only D2 changes in going from FF (RESET) to FB (SPEAK), all other lines can be held at logic '1' for seven ROMCLK cycles before going to logic '0'. The result is three RESETS followed by a SPEAK command. Clocking the counter once more results in Q4 output going high. This signal is fed back to the 4025 NOR gate and prevents any more ROMCLK signals passing to the counter.

On receiving the SPEAK command the synthesiser will commence to pulse the MO line in order to strobe data into the ADD8/DATA port. The MO signal is used to increment a 12 bit counter, (formed from the 4520). This generates the addresses for the EPROM.

The EPROM has speech data programmed into it in such a way that each of the 8 output lines contains a serial data stream. Pressing one of the switches will connect a data output line from the EPROM to the input of the synthesiser. The data stream will therefore pass to the synthesiser as the address counter increments and continues to do so until a stop-code is encountered in the data. The MO signal will cease, the counter will stop and no further data will pass to the synthesiser. Releasing the button will now disconnect the data line and power down the device.

The speech output from the TMS5220 synthesiser is passed through a high pass passive filter to remove any DC signal component. The signal then passes to a low pass second order active filter with a 5 KHz cut-off. The output from the LM308 operational amplifier of the filter is passed via a volume control (internal preset) and AC coupling, to the LM386 audio amplifier. The LM386 drives the internally mounted loudspeaker.

The total peak current drawn from the batteries during a speak phase is 55 m amps. This will give approximately 10 hours continuous use from the PP3 alkaline batteries.

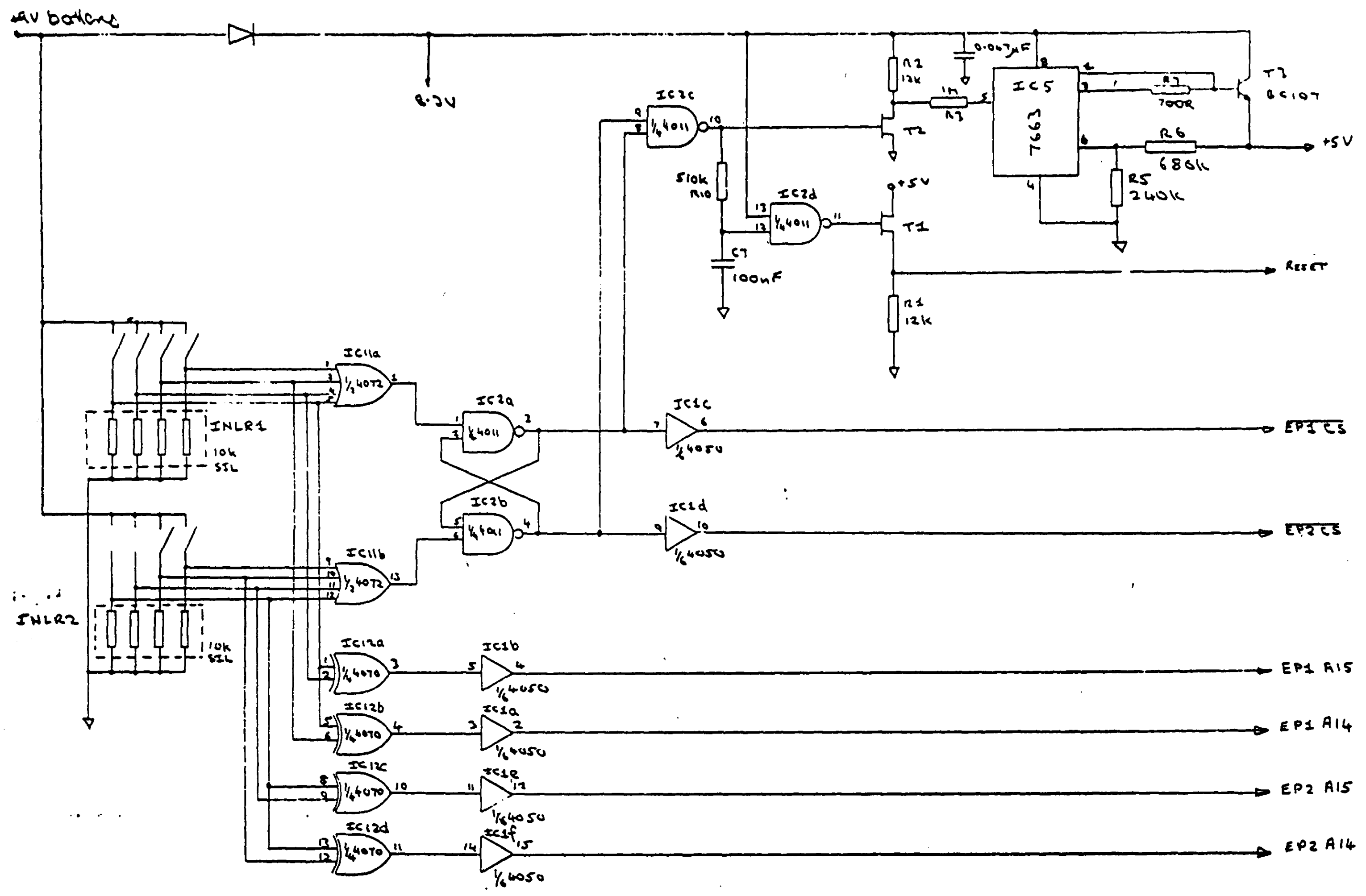

FIGURE A3.1 The Pocket Speech Aid Mk III: Front End Circuit

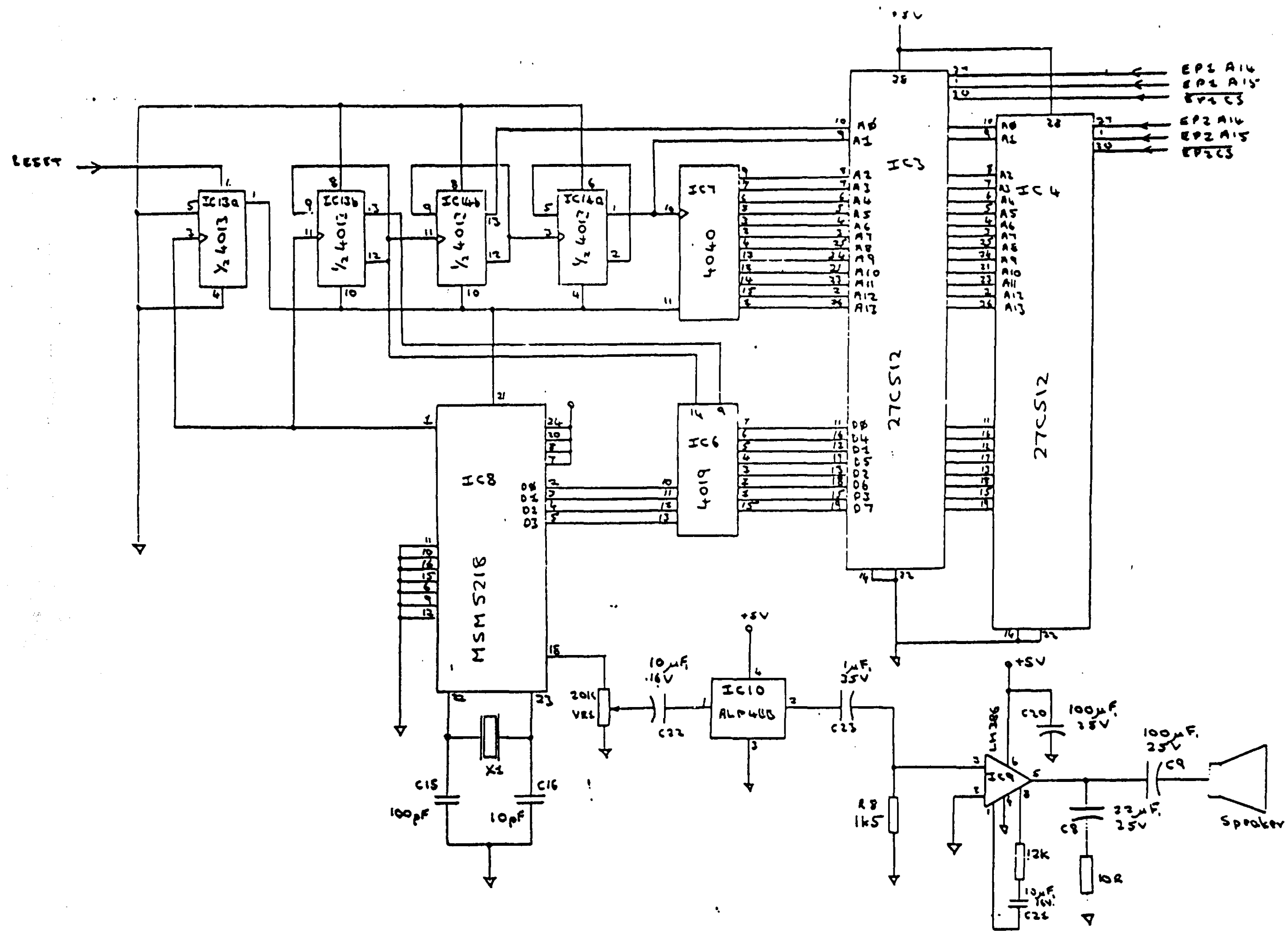

FIGURE A3.2 The Pocket Speech Aid Mk III: Speech Processing Circuit

Circuit Operation

To allow extended battery life the circuit is divided into two sections. One section remains permanently powered, provided the unit is connected to a battery, while the other is only switched on when a key is pressed. The latter will remain on as long as a key is held down, the power being removed when the key is released.

For ease of explanation the permanently powered section will be referred to as the FRONT END (Figure A3.1) and the other section as the SPEECH PROCESSING SECTION (Figure A3.2). The FRONT END is powered directly from a 9V battery. It provides a user interface through two 4 x 1 keypads. It is designed to contain the minimum number of components, while providing an efficient and effective link between the keypads and the speech processing section. The component count is small keeping the power consumption to 10 uA when no keys are pressed. The SPEECH PROCESSING SECTION consists of the components neccessary to retrieve the EPROM resident speech data, convert it to analogue form and output it through a loudspeaker. It is powered from a +5V source, derived from the 9V battery supply through a regulator chip which has logic controlled shutdown capability.

All but one of the signals emanating from the front end are linked through a buffer/level translation chip, 4050 (IC1). This chip is powered from the 5V supply and its outputs remain insensitive to the input signals until a key is pressed. The other remaining signal is coupled through a source follower configuration where the drain on the transistor is connected to +5V.

The Front End

The front end interfaces the two 4x1 keypads to the speech processing section. It detects which key, within which keypad, has been pressed and provides a set of signals to enable the speech processing section to determine which segment of data to retrieve to output the phrase corresponing to that key. Each keypad is assigned an EPROM. Asssociated with each key, within a particular keypad, is a unique 16K segment of data memory within the corresponding EPROM.

On this basis, the four inputs from each keypad are fed into two four input OR gates (IC11) to provide the CS signals for the corresponding EPROMS. To prevent the simultaneous occurence of both CS signals a cross-coupled latch configuration of two two input NAND gates is employed (IC2a , IC2b). The assignment of a key to a 16K segment of memory requires the keys, within each keypad, to generate address lines A15 and A14 for the corresponding 64Kbyte EPROM. This encoding is achieved through the use of 2-input EXCLUSIVE OR gates (IC12). For each keypad two such gates are used. The truth table for the encoding of a set of keys for a keypad is as shown below:

KEY	A15	A14
1	0	0
2	0	1
3	1	0
4	1	1

Since the occurrence of any active low CS signal indicates key depression, they are used to provide a power up signal for the regulator IC (IC5) through a single NAND gate (IC2c) and a common drain configuration. A delayed version of the signal emanating from the same NAND gate is used to provide a power up reset for the counters and latches within the speech processing section.

The regulator IC can only provide 20mA, while 30 - 40 mA is required for some voices. An external transistor, BC107 is used to boost the output current. The advantage of this configuration is that greater output currents are achieved for very little regulator output.

The Speech Processing Section

This section is enabled when a key is pressed and it remains active for the duration a key is held down. The ACPCM speech chip, MSM5218 (ICB), is configured to process 4 bit nibbles at a frequency of 8 KHz. This implies that a 16K byte segment of memory is required for a 4 second phrase. Two flip-flops (IC14a, IC14b) configured as divide-by-2 circuits in conjunction with a 4040 (IC7) are used to provide a 14 bit address counter. This address counter is clocked by a 4KHz signal, derived from a 8 KHz reference from the MSM5218. Initially, the address counter, the speech chip, and the divide-by-2 circuit (IC13b) are held in the reset mode on power up of this section. When the reset is removed the address counter provides sequential addresses required to to fetch the speech data corresponding to the selected phrase. The data output by the EPROM is in byte form. The 4019 (IC6) splits the 8 bit data into 4 bit nibbles, as required by the MSM5218 during the 2 phases of the 4 KHz signal. The data presented to the speech IC is converted to analogue form through its internal D/A convertor. This signal is then presented to a low pass filter with a cut-off frequency of 4 KHz. The final stage in the conversion process consists of a single chip audio amplifier IC, LM386 (IC9) which drives the loudspeaker.

Circuit Description

To extend battery life the unit has been designed so that only a small number of components are permanently powered. These components form the switch sensing part of the circuit and this part is called THE FRONT END in the circuit diagram. The rest of the circuitry is powered only when speech output is required and can be thought of as two separate sections, THE SCANNING SECTION, responsible for phrase selection and THE SPEECH PROCESSING SECTION, responsible for converting the data to a suitable form for the loudspeaker.

The unit has an internal facility for altering the available features. Hence it can be configured, to a certain degree, in accordance with the requirements of the user.

The facilities available are:

1) The number of scans is jumper selectable. A maximum of eight is allowed.
2) The number of phrases is jumper selectable. A maximum of eight is allowed.
3) Upon initiation of the scanning sequence a period of time elapses before the first LED comes on. This is to allow the user time to get ready for the scan. This time can be adjusted internally up to a maximum of 5 seconds.
4) The period of time for which a LED stays lit can also be adjusted internally. The maximimum time allocated for each LED is 5 seconds.

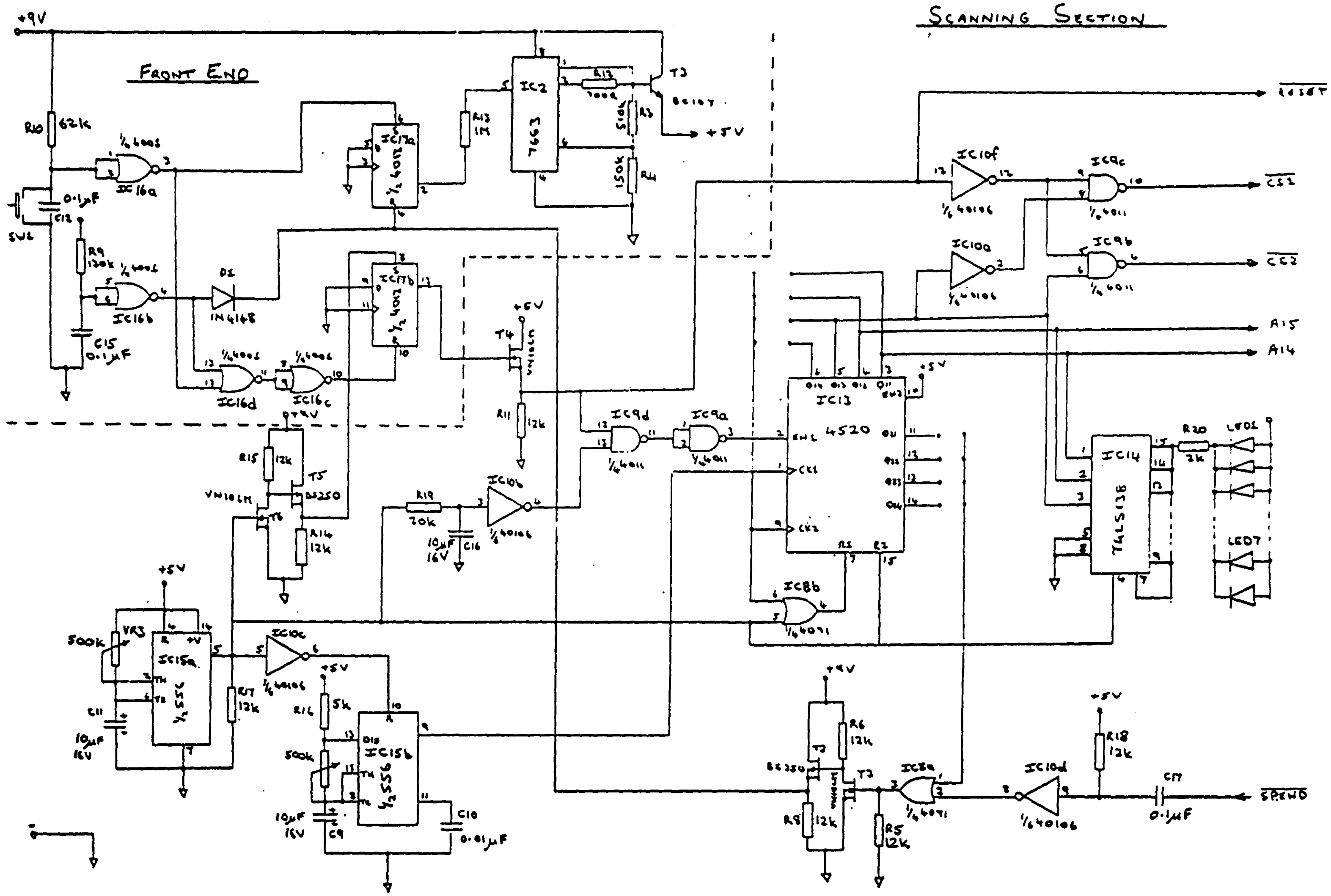

FIGURE A4.1 The Scanning Pocket Speech Aid: Front End Circuit

The Front End (Figure A4.1)

The front end allows user input through a single switch. Its basic function is to detect switch depressions at pre-defined times and then supply the other sections with power and a signal indicating that a switch has been pressed. The main components of this section are two D-type flip-flops connected as SET-RESET latches, latch 1 (IC17a) and latch 2 (IC17b). The function of latch 1 is to either enable or disable the regulator IC 7663 (IC2). If the switch is pressed after the completion of a scanning or a speaking period latch 1 will enable the 7663 to initiate another scan. If on the other hand the unit has just completed a scanning/speaking period the latch will disable the 7663 in accordance with a power-down signal received from any of the other sections. The reset input of latch 1 receives signals from three sources. One from the scanning section to indicate that no phrase selection has taken place. Another is from the speech processing section to signal the completion of a spoken phrase. The third is from a power up circuit from within the front end, simply to ensure that latch 1 resets on battery connection. Latch 2 is used to detect the second allowed switch depression, which indicates that the user wants to select a phrase.

On the first switch depression the Q output of latch 2 is forced high through a signal received from the power-up monostable (IC15a) in the scanning section. This renders latch 2 insensitive to switch depressions until the output of the monostable goes low. It is then ready to accept switch inputs. The Q output of this latch serves as a number of functions in the other sections. It is used to hold the counters and latches in the speech processing section in the reset mode during a scanning sequence and to remove the reset if phrase selection takes place. It holds the phrase counter (IC13a) in the selected state and strobes its outputs to the speech processing section so that the corresponding phrase can be output.

The Scanning Section (Figure A4.1)

This section provides the speech processing section with a set of signals which act as pointers to segments of speech data stored in the EPROMS. These pointers are merely the states of the outputs of a counter, previously referred to as the phrase pointer counter. The set of signals sent are those which allow the selection of one of two 64K byte EPROMS, and a 16K byte segment within the selected EPROM. These signals are the chip select signals CS1,CS2 and the address lines A15 and A16. A visible indication of the present state of the phrase pointer counter during a scan is provided by a lit LED. The lit LED indicates which phrase is currently being pointed to during a scan.

Another counter (IC13b) is used to provide a power-down signal to latch 1 in THE FRONT END when no phrase selection takes place. This counter is configured to count the scan cycles of the phrase pointer counter and its outputs are connected to a jumper (J1) which allows any one of its outputs to be used as the power-down signal. The number of scan cycles can be set to 1,2,4 or 8.

On power up the monostable goes high and remains so for a period of time determined by C11 and VR3. This output is used to initialise the counters and to disable the outputs of the multiplexer, the 74LS138 (IC14) and the astable multivibrator (IC15b). When the monostable output goes low the astable becomes active and provides the clock pulses for the phrase pointer counter. One drawback of the astable circuit is

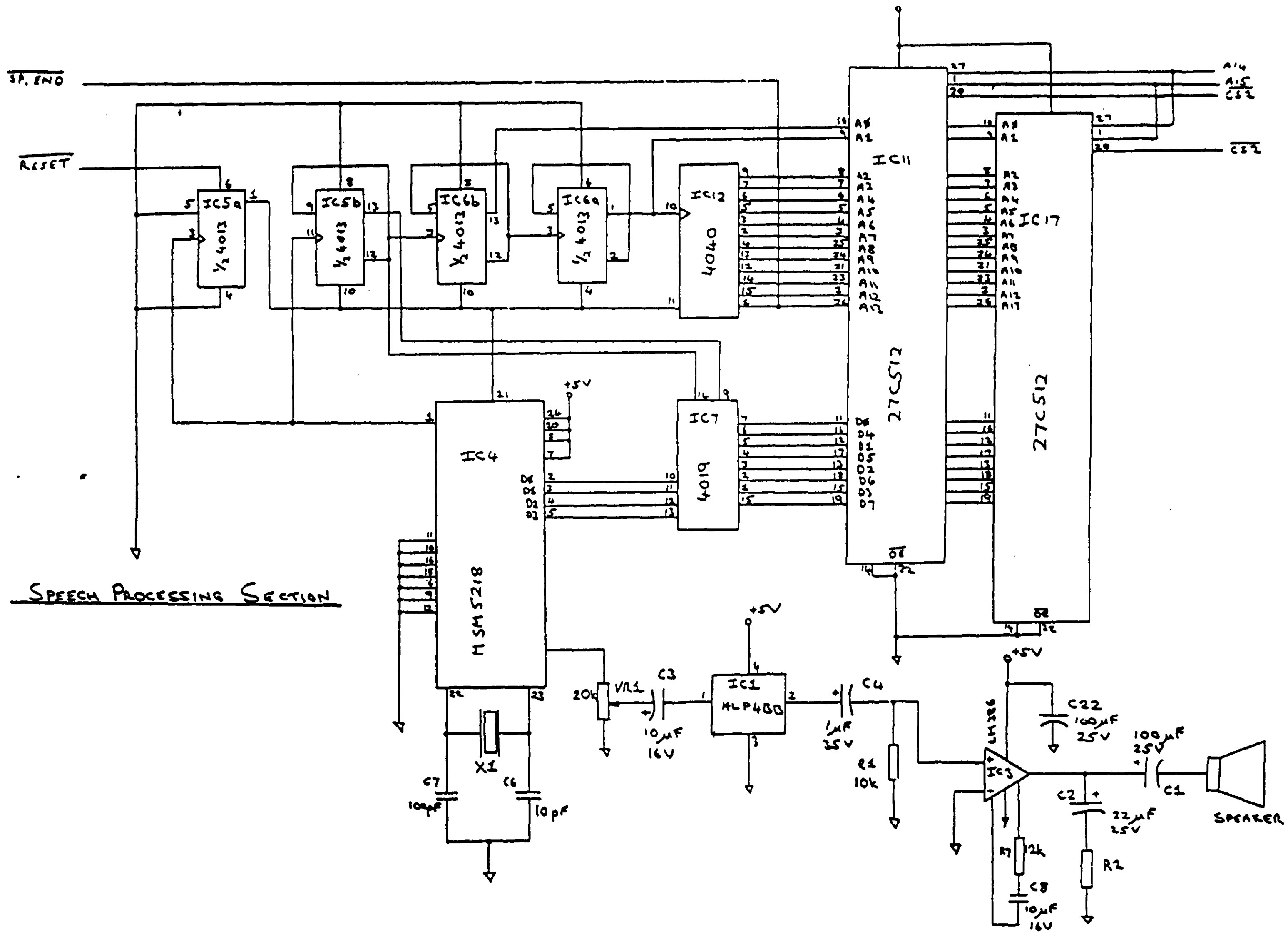

FIGURE A4.2 The Scanning Pocket Speech Aid: Speech Processing Circuit that as soon as the monostable output goes low the astable output immediately goes high. Therefore a delayed version of the monostable output is used to inhibit the clock input of the phrase pointer counter during the first low to high transition of the astable.

The Speech Processing Section (Figure A4.2)

This section outputs a phrase in accordance with the signals received from the scanning section. The ADPCM speech IC, MSM5218 (IC4), is configured to process 4-bit nibbles at a frequency of 8 KHz. This implies that a 16 Kbyte segment of memory is required for a 4 second phrase.

Two 64 Kbyte EPROMS. 27C512s (IC11, IC17) are used to give a data capacity of 128 Kbytes and hence a maximum speech capability of 8 X 4 second phrases. Two flip-flops (IC6) configured as divide-by-two circuits in conjunction with a 4040 (IC12) are used to provide a 14 bit address counter. This address counter is clocked by a 4 KHz signal, derived from another divide-by-two counter (IC15b) from an 8 KHz reference from the MSM5218.

The address counter, the speech IC and the divide-by-two circuits are held in the reset mode during the scanning sequence. If a phrase selection takes place, the reset is removed and the address counter provides the sequential addresses required to fetch the speech data corresponding to the selected phrase. The flip-flop of IC15a serves to synchronise the address counter on removal of the reset.

The data output by the EPROM is in byte form and the 4019 (IC7) is used to split the 8 bit data into 4 bit nibbles, as required by the MSM5218, during the two phases of the 4 KHz clock cycle. The data presented to the speech IC is converted to analogue form by its internal D/A converter. The signal is then presented to a low pass filter with a cut-off frequency of 4 KHz. The final stage in the conversion process consists of a single chip audio amplifier IC, LM386 (IC3), which drives a speaker.

APPENDIX 5

AN INSTRUCTION BOOKLET TO ASSIST THE USER OF A CONVERSATION PROMPTER

About conversation prompting

Conversation prompting is a means of communication between two or more people, one of whom cannot speak and has to use the speech aid device known as the Conversation Prompter. First of all take a look at your device. Press the buttons one at a time and listen to the phrases. Notice that the button can be released before the end of the phrase thereby providing a different word or phrase. The phrases and their explanations are listed at the end of this book.

If you are having difficulty in pressing the buttons your speech therapist may be able to help by providing an attachment to clip over the conversation prompter box. Please make sure she is aware of any problems you are having.

There is much more to conversation prompting however than simply pressing the buttons on the prompter. The skill which you are trying to acquire is that of steering the direction of a conversation by using the few phrases available to you. You can let the friend who is conversing with you know that you would like him to start speaking, keep going on the same subject, change to a different subject, explain himself better, etc. You must however make use of every facility available to you, not just the conversation prompter. You may attempt to get messages over by making signs, gestures, etc, in addition to using the phrases. There are several simple rules which can help you to make good use of your conversation prompter and the following lessons are designed to teach you these rules.

Part 1

Hello Dad, are you feeling any better?

YES

That's good. Still getting the headaches?

NO

That's good - I think you are on the mend

Your friend will be instructed to phrase questions in such a way that you can answer Yes or No. Note that you have to take your finger off the key after "Yes" has been said or else the full phrase "Yes, keep talking to me" will be spoken by the device.

Likewise you have to take your finger off after pressing "No" otherwise the device will say "No, I prefer something else". Sometimes you will want the full phrases to be spoken but on most occasions simply Yes or No will be sufficient.

Part 2

Do you want me to stay and talk for 10 minutes?

WHAT A GOOD IDEA

Notice that in this case the phrase "What a good idea" was used rather than simply Yes. This is a much more positive phrase and indicates that you are very enthusiastic about the idea. Yes would indicate agreement but would not have the same degree of enthusiasm associated with it.

If you are really keen on the suggestion made by your friend, use the phrase "What a good idea".

Part 3

We took the boys to the gang show yesterday. They really enjoyed it but I'm afraid Margaret and I thought it was pretty dreadful. It is not the kids' faults - they do their best - its just that the script is so corny.

Mind you there was one really good act. It was a band all playing synthesisers. They ....

HOLD ON. WHAT

Synthesisers. That's these electronic piano things.

Your friend will be encouraged to do most of the talking. After all if he is not speaking you cannot change the direction of the conversation.

Notice that two keys were used in this case, "Hold on" followed by "What". "Hold on" is a very useful key as it stops your friend from speaking so that he is listening clearly for your next instruction. It also gives you time to find the appropriate key to get your message over.

Part 4

Well the wee girl in number 45 was playing one of them.

You know who I mean.

NO

You know - the wee red haired girl with glasses.

Walks ....

YES - KEEP ON TALKING

Your friend will normally go through several suggestions until he gets his message over. As soon as you know the answer use the phrase "Yes - keep talking". This lets your friend know that you have got the message and he can carry on. Otherwise he will keep prompting you with more clues.

Part 5

Well, they played some really good music - all Lennon & McCartney tunes. The first .....

HOLD ON - COULD WE CHANGE THE SUBJECT

Sorry, I forgot you didn't like that kind of music.

Don't be afraid to ask your friend to change the subject. Part of the knack of having a successful conversation between two people using the conversation prompting technique is for the friend to keep talking on a subject until you indicate that he should change to something else. You are not being impolite; he will recognise that this is simply part of the conversation method.

Part 6

I was going to tell you about the football yesterday - are you interested?

YES, WHAT A GOOD IDEA

Aberdeen won 3-1. They played a lot better this week. Jim Bett was back in the team and that seemed to make a big difference.

And Rangers won 5-1 against Dundee. Celtic .....

HOLD ON. WHO

Who - Oh! who scored. Davy Cooper scored 3. I can't remember the rest.

After pressing "Who" in the above example the user of the conversation prompter made a sign to his friend by giving a small kick with his foot and put his hands up in the air to imitate the scorer of a goal. Making signs is something which you will have to do frequently. The conversation prompter gets the sentence going and gives your friend a lead as to what you are signing about. You will be surprised how far a simple word or phrase followed by a gesture can go when it comes to getting messages over. This is an extremely important part of conversation prompting.

Part 7

Celtic ....

HOLD ON

What does this say. "Could you fix my toaster". Is it not working?

NO

OK. I'll take it away with me.

In this example the user of the device has made a note for his son before he arrived. He suddenly remembered this and, in case he forgot, interrupted his son again and drew his attention to the note. Here again he is making use of another means of communication. Passing written mesages back and forward is not a terribly immediate form of coversation but it is very useful for getting messages over which can be prepared in advance.

Part 8

I think I'd better be going now.

WHERE ARE YOU GOING?

I'm going to Motherwell. John has a badminton match, and I've to run the team over. Peter will go and bring them back at 5 o'clock.

I'll be off then.

WHEN WILL I SEE YOU AGAIN?

I'll look in tomorrow evening. Oh no - I won't. I'm going down to Greenock.

WHY?

We've been invited to friends for dinner.

WHO?

Tom and Fiona - they used to stay down the road from us.

YES

OK. Must go. Bye for .....

HOLD ON

You want me to stay.

NO. I'D PREFER SOMETHING ELSE

Oh, the toaster. Sorry I forgot. Is it in the kitchen? OK, I'll get it. There we are. Bye for now.

Note from the above passage how you have the ability to ask lots of questions.

This conversation has demonstrated to you some of the techniques which can lead to efficient use of the conversation prompter. Like all forms of communication the technique has to be worked at and perfected and this takes time and practice. You will no doubt make mistakes at first. Your friends will also take some time to get used to the best way of having a conversation with you. You must take every opportunity to practice together and you will find that gradually your familiarity with the technique will improve and your conversations will become more fluid. Given time you should find the conversation prompter a fairly efficient and worthwhile aid to communication.

APPENDIX 6 - CASE STUDIES

A communication aid case study form was developed and is in the form of a 'mini' communication profile that is manageable and provides feedback about various aspects of the devices.

The form contains personal details, medical diagnosis and the physical and mental disability of the patient. There is space for environmental factors, communication assessment prior to the introduction of an aid and communication assessment with the aid. Several features are scored on a scale from '*' to '*****'.

Overall Rating:	Takes into account all of the features stated below and rates the device in terms of its usefulness.
Ease of Use:	Refers to the patient's physical and mental ability to operate the device.
Sound Quality:	Refers to the standard of speech from the patient and the listeners point of view.
Portability:	The device is scored in terms of ease of transportation for a particular individual.
Encouragement:	Refers to the encouragement given by the patient's helpers.
Acceptance:	Refers to the patient's acceptance of a need for the synthetic speech aid.
Vocabulary range:	Scores the vocabulary range for a device for a particular patient's requirements.

The forms were filled in by the speech therapist after consultation with the patients and their helpers.

COMMUNICATION AID CASE STUDY

NAME: P.K. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 1920

MEDICAL DIAGNOSIS: Cerobrovascular accident

PHYSICAL DISABILITIES: Right Hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dyspraxia

ENVIRONMENTAL FACTORS

The patient is unable to go out and is permanently housebound. She has few visitors and has become very passive. Her sister looks after her and anticipates all her needs.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She uses gesturing and facial expression with some limited stereo-typed language. She is unable to write and relies heavily on her sister.

COMMUNICATION ASSESSMENT - WITH AID

There was no need for the communication aid in this situation as passivity had been induced and her sister anticipated all her needs.

DEVICE RATING:

Portability is not an important feature as the patient spends almost all of her time sitting in a chair. She did not receive any encouragement from her sister. Therapy sessions produced some language recovery which allowed the expected extension of social life but there was no real need for a communication aid.

OVERALL RATING: *
EASE OF USE: ****
SOUND QUALITY: *****
PORTABILITY: ***
ENCOURAGEMENT: *
ACCEPTANCE: *
VOCABULARY RANGE: ****

NAME: M.S. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Right Hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Unable to express herself

ENVIRONMENTAL FACTORS

The patient is in a geriatric hospital and is unable to communicate with her visitors or with nursing staff.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She has adequate comprehension skills but is unable to express herself.

COMMUNICATION ASSESSMENT - WITH AID

The patient was enthusiastic about using a conversation prompter and initially found it useful. She eventually rejected the device because of poor speech quality, insufficient therapy (circumstantial) and insufficient encouragement from her carers. Her device had female LPC speech and she had heard the superior quality of the ADPCM device.

DEVICE RATING

The overall rating is moderate. She is comfortable with the conversation prompting technique but was unhappy with the quality of speech which was inferior to the usual standard produced for this device. She has recently been issued with the new ADPCM device.

OVERALL RATING: ***
EASE OF USE: ****
SOUND QUALITY: *
PORTABILITY: *****
ENCOURAGEMENT: ***
ACCEPTANCE: **
VOCABULARY RANGE: ****

NAME: M.S. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Limb weakness

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysphonia, poor volume

ENVIRONMENTAL FACTORS

This patient lives in a Christian Community Home and has many visitors.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her communication is normal for part of the day but she is only able to speak for a very short period of time before becoming Aphonic.

COMMUNICATION ASSESSMENT - WITH AID

The communication aid gave her the ability to communicate at all times. The device suited her better than an amplifier as it compensated for her fatigue problem. She did not have to make as much effort to speak.

DEVICE RATING

The overall rating is moderate. The patient found the device reasonably easy to operate even although she had a degree of limb weakness.

OVERALL RATING:	***
EASE OF USE:	***
SOUND QUALITY:	****
PORTABILITY:	*****
ENCOURAGEMENT:	****
ACCEPTANCE:	****
VOCABULARY RANGE:	****

COMMUNICATION AID CASE STUDY

NAME: M.M. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: Extremely depressed

STATE OF SPEECH: None

ENVIRONMENTAL FACTORS

The patient lives at home with her family.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patients only method of communication was by gesturing.

COMMUNICATION ASSESSMENT - WITH AID

She did not use the aid as she preferred gesturing. She did not have the motivation to use the conversation prompter.

DEVICE RATING

The overall rating is low. The patient is depressed and unable to motivate herself to use the speech aid.

OVERALL RATING:	*
EASE OF USE:	*
SOUND QUALITY:	****
PORTABILITY:	****
ENCOURAGEMENT:	*
ACCEPTANCE:	*
VOCABULARY RANGE:	*

NAME: M.T. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 1934

MEDICAL DIAGNOSIS: CVA

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Fluent dysphasia

ENVIRONMENTAL FACTORS

The patient leads a normal active life and requires to communicate in a wide variety of situations.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient still had some functional speech but his speech was interspersed with jargon.

COMMUNICATION ASSESSMENT - WITH AID

The speech aid gave the patient good psychological support. He knew that he could revert to the speech aid if his own speech was difficult. He has experienced some language recovery. The device was used only occasionally but it helped him through a difficult period.

DEVICE RATING

The patient was enthusiastic about the idea of conversation prompting and was very keen on using the device. His language recovery meant that the device was rarely used but provided him with support.

OVERALL RATING: ***
EASE OF USE: ***
SOUND QUALITY: *****
PORTABILITY: *****
ENCOURAGEMENT: ****
ACCEPTANCE: *****
VOCABULARY RANGE: ****

NAME: E.W. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Right Hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dyspraxia

ENVIRONMENTAL FACTORS

The patient was facing an imminent change of environment from hospital to a geriatric home. She was highly motivated to enjoy her new home and had a great desire to communicate.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She had tried using a Possum 16 for basic needs but she found the device too slow and required much more than basic needs. The patient has good comprehension abilities. She is limited to yes/no responses and is unable to use phrases or initiate conversations. Her gesture and facial expression skills are limited.

COMMUNICATION ASSESSMENT - WITH AID

She is very good at using the technique of conversation prompting. She is able to initiate, hold attention and keep a conversation flowing. The device can give her good control of her environment.

DEVICE RATING

The overall communication rating for the device is high. The patient is highly motivated to communicate. She enjoys using the speech aid and is very happy with all of its features.

OVERALL RATING: *****
EASE OF USE: *****
SOUND QUALITY: *****
PORTABILITY: *****
ENCOURAGEMENT: ***
ACCEPTANCE: *****
VOCABULARY RANGE: ****

COMMUNICATION AID CASE STUDY

NAME: M.J. **DEVICE:** Pocket Speech Aid (Conversation prompter)

DATE OF BIRTH: 1921

MEDICAL DIAGNOSIS: Unknown

PHYSICAL DISABILITIES: Poor hand and limb movement

MENTAL DISABILITIES: None

STATE OF SPEECH: None

ENVIRONMENTAL FACTORS

The patient has suffered many years of disability and the routine environment has introduced a degree of passivity. He spends his time in hospital or in sheltered housing.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient uses gesturing, facial expression and the Texas Instruments speak 'n spell toy.

COMMUNICATION ASSESSMENT - WITH AID

He used the conversation prompter in hospital but preferred the unlimited vocabulary of the speak 'n spell device when he was at home.

DEVICE RATING

In the hospital situation the overall communication rating is high as the device enabled him to request information and assistance very quickly. Most of the device features were acceptable to this patient.

OVERALL RATING:	*****
EASE OF USE:	***
SOUND QUALITY:	*****
PORTABILITY:	****
ENCOURAGEMENT:	*****
ACCEPTANCE:	*****
VOCABULARY RANGE:	***

NAME: D.M. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 31/7/19

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient has not accepted the prognosis.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient has good mobility and is able to use all fingers. He gestures and uses writing as his preferred mode of communication.

COMMUNICATION ASSESSMENT - WITH AID

He does not use the device as he has not accepted the need for an aid. He has not accepted the prognosis and it is possible that it is the wrong time to introduce an aid. The patient is distressed, anxious and pre-occupied with his dysphagia problems.

DEVICE RATING

The overall communication rating at this time is low and most of the device features are not relevant.

OVERALL RATING:	*
EASE OF USE:	**
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	*
ACCEPTANCE:	*
VOCABULARY RANGE:	*

NAME: J.B. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 1920

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Right Hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dyspraxia, Dysphasia

ENVIRONMENTAL FACTORS

The patient spends her time at home and in hospital. The routine environment has induced passivity but she has become frustrated by her inability to communicate with her Home Help and her daughter.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The preferred mode of communication was unintelligible words and facial expression. She was unable to provide information and her dyspraxia caused her to mix up yes/no responses.

COMMUNICATION ASSESSMENT - WITH AID

The device helped her to monitor her yes/no responses and was useful for a few weeks. The patient finally rejected the aid preferring her old methods of communication. Statements and questions related to specific situations were useful until listeners adjusted and anticipated her needs.

DEVICE RATING

Her daughter was enthusiastic about the device as it helped her to understand her mothers needs. The device was useful as a monitoring device for the patients own speech.

OVERALL RATING:	****
EASE OF USE:	***
SOUND QUALITY:	*****
PORTABILITY:	**
ENCOURAGEMENT:	*****
ACCEPTANCE:	*****
VOCABULARY RANGE:	*****

NAME: M.M. **DEVICE:** Pocket Speech Aid (Conversation prompter)

DATE OF BIRTH: 1935

MEDICAL DIAGNOSIS: Parkinson's disease

PHYSICAL DISABILITIES: Severe intermittent tremor

MENTAL DISABILITIES: None

STATE OF SPEECH: Intermittent dysarthria

ENVIRONMENTAL FACTORS

The patient is housebound and has frequent stays in hospital. She needs to communicate with her family and the nursing staff. She is extremely frustrated at her inability to communicate during periods of tremor.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She is in the the tremor state two or three times a day and each period can last for up to 2 hours. She is unable to communicate at all during these spells.

COMMUNICATION ASSESSMENT - WITH AID

Under normal conditions the patient is able to operate the device but during the tremor period it is physically too difficult for her to operate. It is possible that she might be able to operate the single switch scanning speech aid.

DEVICE RATING

The overall communication rating is low as the patient cannot cope with the device during tremor.

OVERALL RATING:	*
EASE OF USE:	*
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	*****
ACCEPTANCE:	**
VOCABULARY RANGE:	**

COMMUNICATION AID CASE STUDY

NAME: M.D. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 1920

MEDICAL DIAGNOSIS: Pseudo-bulbar Palsy

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: Emotionally labile

STATE OF SPEECH: Non vocal

ENVIRONMENTAL FACTORS

The patient spends almost all of his time at home with his wife. He has few communication needs as he is unable to leave the house and has few other acquaintances.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

His preferred mode of communication is gesturing. He is unable to write but he occasionally uses a Canon Communicator for single words.

COMMUNICATION ASSESSMENT - WITH AID

The patient is an expert gesturer and his wife is able to anticipate his every need. The speech aid serves no useful purpose in close relationships as language adjustments have been made by listeners.

DEVICE RATING

The overall communication rating is low as the patient preferred his existing methods of communication.

OVERALL RATING:	*
EASE OF USE:	*****
SOUND QUALITY:	*****
PORTABILITY:	****
ENCOURAGEMENT:	*
ACCEPTANCE:	***
VOCABULARY RANGE:	*

COMMUNICATION AID CASE STUDY

NAME: A.M. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

The patient is in long term hospital care. She is frustrated by her inability to initiate conversations in the ward.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her preferred mode of communication is writing which is useful for providing information but it is difficult for her to gain attention.

COMMUNICATION ASSESSMENT - WITH AID

The prompter enables her to gain and hold attention. She uses the device in conjunction with writing and finds it useful in encouraging other patients to talk. She prefers to write when providing information and uses the speech aid for requesting information.

DEVICE RATING

The patient had 4 therapy sessions on the techniques of conversation prompting. She is able to use the prompter in conjunction with writing to great effect. The overall communication rating is very high.

OVERALL RATING:	*****
EASE OF USE:	*****
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	***
ACCEPTANCE:	*****
VOCABULARY RANGE:	*****

NAME: J.K. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 1913

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Poor vision , below knee amputation

MENTAL DISABILITIES: None

STATE OF SPEECH: Severe dyspraxia , moderate dysphasia

ENVIRONMENTAL FACTORS

The patient is unable to go out and is permanently housebound. Her communication needs are limited to aspects of daily living and her main communication is with her sister.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She uses gesturing, pointing and facial expression. She is unable to write due to her poor vision.

COMMUNICATION ASSESSMENT - WITH AID

Gesturing and pointing remain her preferred mode of communication. She is unable to realise the potential of a synthetic speech aid.

DEVICE RATING

The patient received 2 therapy sessions on the use of the device and received limited ensouragement from her family. She has difficulty operating the device due to her poor vision.

OVERALL RATING:	*
EASE OF USE:	**
SOUND QUALITY:	****
PORTABILITY:	*****
ENCOURAGEMENT:	****
ACCEPTANCE:	****
VOCABULARY RANGE:	***

NAME: G.B. **DEVICE:** Pocket Speech Aid (Conversation Prompter)

DATE OF BIRTH: 7/8/13

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Right hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Fluent dysphasia

ENVIRONMENTAL FACTORS

The patient spends all of his time in hospital and requires to communicate with family, friends and nursing staff. He has many and varied interests i.e. football, music, medicine and reading.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He is a poor communicator and has a tendency to be verbose. His main mode of communicatior is speech. His main difficulty is in monitoring his own speech which is fluent.

COMMUNICATION ASSESSMENT - WITH AID

The patient is an intelligent man who is socially competent and is able to use the conversation prompter as a controller for his own speech.

DEVICE RATING

The overall communication rating is moderate but could be higher if the vocabulary was personalised for this particular patient.

OVERALL RATING: ***
EASE OF USE: ****
SOUND QUALITY: *****
PORTABILITY: *****
ENCOURAGEMENT: *****
ACCEPTANCE: *****
VOCABULARY RANGE: **

COMMUNICATION AID CASE STUDY

NAME: L.M. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 8/7/22

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: No physical deterioration

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria, verbal expression unintelligible

ENVIRONMENTAL FACTORS

The patient goes on shopping trips and visits her sons and daughters but she does spend most of her time at home. She needs to communicate with her spouse, children and visitors to the house.

Encouragement is provided by her family but she is frustrated by her inability to initiate and maintain conversation.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient communicates through writing, gesturing and with help from her husband. Speech is very poor and reduced to basic vowel phonemes such as /a/ and /ai/.

COMMUNICATION ASSESSMENT - WITH AID

She had two therapy sessions on the use of the aid. The device allowed her to use the telephone with friends and family. The device allowed her to request assistance but did not enable her to provide or emotional expression.

DEVICE RATING

The patient found the device easy to operate. Portability was an important feature enabling her to use the telephone at home, in hospital and in public places. She received a great deal of encouragement and found the aid acceptable to listeners.

OVERALL RATING:	*****
EASE OF USE:	*****
SOUND QUALITY:	****
PORTABILITY:	*****
ENCOURAGEMENT:	*****
ACCEPTANCE:	*****
VOCABULARY RANGE:	***

COMMUNICATION AID CASE STUDY

NAME: M.T. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1934

MEDICAL DIAGNOSIS: CVA

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Fluent dysphasia

ENVIRONMENTAL FACTORS

This patient leads a normal active life and requires to communicate in a wide variety of situations.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient had some functional speech but his speech was interspersed with jargon.

COMMUNICATION ASSESSMENT - WITH AID

The aid was used to enable him to use the telephone normally. He was unable to use the telephone without the support of the aid.

DEVICE RATING

The overall rating for this device is high because it enabled the patient to communicate over the telephone which he would otherwise not attempt to use. He still uses his own speech but he occasionally uses the speech aid when he runs into difficulties. The aid is used to give him confidence in himself.

OVERALL RATING:	****
EASE OF USE:	*****
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	***
ACCEPTANCE:	****
VOCABULARY RANGE:	****

COMMUNICATION AID CASE STUDY

NAME: M.M. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1916

MEDICAL DIAGNOSIS: Pseudobulbar Palsy

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysphasia , word finding difficulty.

ENVIRONMENTAL FACTORS

The patient lives alone and needs to communicate with her neighbour and her son. She is permanently housebound and requires assistance for her everyday needs.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She is able to write and relies on visits from her neighbour and son.

COMMUNICATION ASSESSMENT - WITH AID

She uses the telephone aid to contact her son and neighbour.

DEVICE RATING

The device enabled the patient to contact people for assistance without having to wait for a visit.

OVERALL RATING:	****
EASE OF USE:	****
SOUND QUALITY:	***
PORTABILITY:	*****
ENCOURAGEMENT:	***
ACCEPTANCE:	*****
VOCABULARY RANGE:	****

COMMUNICATION AID CASE STUDY

NAME: M.A. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1937

MEDICAL DIAGNOSIS: Funtional Aphonia

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: None

ENVIRONMENTAL FACTORS

This patient leads a normal active life. She lives alone and would normally communicate regularly with her family on the telephone. She would like to continue her social life as much as possible.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her social life had been disrupted. Her communication on the telephone was limited to tapping noises for yes/no responses to questions.

COMMUNICATION ASSESSMENT - WITH AID

She was able to communicate with her family and friends on the telephone.

DEVICE RATING

The overall rating for the device is high as it helped her to continue her social life and acted as a therepeutic tool allowing her to relax and consider her problems. The personalised vocabulary was designed by herself and her son. She was very enthusiastic about using the device and received a lot of encouragement from her family.

OVERALL RATING: ****
EASE OF USE: *****
SOUND QUALITY: ***
PORTABILITY: *****
ENCOURAGEMENT: *****
ACCEPTANCE: ***
VOCABULARY RANGE: ****

COMMUNICATION AID CASE STUDY

NAME: D.M. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 31/7/19

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient has not accepted the prognosis.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient has good mobility and is able to use all fingers. He gestures and uses writing as his preferred mode of communication.

COMMUNICATION ASSESSMENT - WITH AID

He does not use the device as he has not accepted the need for an aid. He has not accepted the prognosis and it is possible that it is the wrong time to introduce an aid.

DEVICE RATING

The overall communication rating at this time is low and most of the device features are not relevant.

OVERALL RATING:	*
EASE OF USE:	**
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	*
ACCEPTANCE:	*
VOCABULARY RANGE:	*

COMMUNICATION AID CASE STUDY

NAME: M.B. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Cerobrovascular accident

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysphasia

ENVIRONMENTAL FACTORS

This patient lives at home with his wife but would like to have a degree of independence. He has difficulty in communicating due to difficulty in accessing his own language.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient was unable to answer the telephone and had to rely on his wife.

COMMUNICATION ASSESSMENT - WITH AID

The patient was able to use the telephone by using the machine to help him access his own language. He could press a button on the device and then repeat the phrase himself on the telephone.

DEVICE RATING

The overall rating for the device is high because it enabled the patient to communicate on the telephone using the device as a prompt for his own speech.

OVERALL RATING:	****
EASE OF USE:	****
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	***
ACCEPTANCE:	****
VOCABULARY RANGE:	****

COMMUNICATION AID CASE STUDY

NAME: J.D. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1943

MEDICAL DIAGNOSIS: Multiple sclerosis

PHYSICAL DISABILITIES: Wheelchair bound

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

The patient spends all of his time at home and depends on his family to meet his needs.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He used gesturing and dysarthric speech as his preferred mode of communication. He was unable to use the telephone.

COMMUNICATION ASSESSMENT - WITH AID

The patient was enthusiastic about using the device as a telephone aid. He found the device physically difficult to operate because of his hand tremor. If a single switch version of the device was available at this time it would have been more successful.

DEVICE RATING

The device helped to stimulate the patients interest in communication and gave him a degree of independence. Single switch operation would be preferred to direct selection.

OVERALL RATING: ****
EASE OF USE: ***
SOUND QUALITY: ****
PORTABILITY: *****
ENCOURAGEMENT: *****
ACCEPTANCE: *****
VOCABULARY RANGE: ****

COMMUNICATION AID CASE STUDY

NAME: D.R. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia , Limb weakness

MENTAL DISABILITIES: None

STATE OF SPEECH: Severe dysarthria

ENVIRONMENTAL FACTORS

The patient spent most of his time at home. He lived with his mildly handicapped son. He was a retired surgeon and was able to accept the prognosis.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He used writing and gesturing as his preferred mode of communication. He was unable to use the telephone.

COMMUNICATION ASSESSMENT - WITH AID

He used the device to answer the telephone but did not attempt to make his own calls.

DEVICE RATING

The patient found the device easy to operate and liked using it. The overall communication rating is moderate as it only slightly increased his communication abilities.

OVERALL RATING:	***
EASE OF USE:	*****
SOUND QUALITY:	****
PORTABILITY:	*****
ENCOURAGEMENT:	***
ACCEPTANCE:	*****
VOCABULARY RANGE:	***

COMMUNICATION AID CASE STUDY

NAME: J.H. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1923

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Right hemiplegia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dyspraxia

ENVIRONMENTAL FACTORS

The patient spends almost all of his time at home. He has reasonable comprehension skills but very little intelligible speech. He is not very mobile and uses a walking stick.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He uses gesturing and facial expression as his preferred mode of communication. He would like to use the telephone.

COMMUNICATION ASSESSMENT - WITH AID

He was initially very enthusiastic about the telephone aid and used it for communicating with his wife and friends. After a period of several weeks he lost interest in the device. His motivation was not strong enough.

DEVICE RATING

The patient did not find the device easy to operate with right hemiplegia and after a short period of time the device was rejected and therefore did not increase his communication ability.

OVERALL RATING:	**
EASE OF USE:	**
SOUND QUALITY:	****
PORTABILITY:	**
ENCOURAGEMENT:	***
ACCEPTANCE:	***
VOCABULARY RANGE:	***

COMMUNICATION AID CASE STUDY

NAME: S.P. **DEVICE:** Pocket Speech Aid (Telephone)

DATE OF BIRTH: 1926

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia, able-bodied

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient lives at home. Her previous occupation was as a clerkess/typist. She has good hand movement and is therefore able to operate keyboard devices efficiently. She leads a normal active life.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her preferred mode of communication is writing and gesturing. She occasionally uses the Votrax, particularly when communicating with her young grand-children.

COMMUNICATION ASSESSMENT - WITH AID

The patient prefers to write and gesture in normal conversation but finds the device useful for telephone communication.

DEVICE RATING

The patient likes the voice quality of the device which is superior to her previous device (Votrax). She enjoys using the Pocket Speech Aid and finds all the features acceptable.

OVERALL RATING: *****
EASE OF USE: *****
SOUND QUALITY: *****
PORTABILITY: *****
ENCOURAGEMENT: *****
ACCEPTANCE: *****
VOCABULARY RANGE: ***

COMMUNICATION AID CASE STUDY

NAME: N.M **DEVICE:** Pocket Speech Aid (car situation)

DATE OF BIRTH: 1915

MEDICAL DIAGNOSIS: Cerobrovascular accident

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: Confused at times

STATE OF SPEECH: Jargon Dysphasia, inappropriate language

ENVIRONMENTAL FACTORS

The patient is frequently driven in a car by volunteers from the Stroke club and needs to be able to give directions.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient gives inapropriate instructions to the volunteers driving the car.

COMMUNICATION ASSESSMENT - WITH AID

The patient persisted in communicating in his normal way and was reluctant to make use of the device.

DEVICE RATING

The overall communication rating is low because the device did not help the patient to clarify the situation. He was physically able to use the device but his mental problem meant the device was not used.

In the therapy session in which the aid was introduced he was able to use the device and felt that the sound quality was very good. This patient requires more therapy and more encouragement from his wife and helpers before an aid of this type could be effective. The vocabulary range covered all the requirements for this situation.

OVERALL RATING:	*
EASE OF USE:	**
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	*
ACCEPTANCE:	*
VOCABULARY RANGE:	*****

NAME: M.F. **DEVICE:** Pocket Speech Aid (car situation)

DATE OF BIRTH: 1916

MEDICAL DIAGNOSIS: Cerobrovascular accident

PHYSICAL DISABILITIES: Wheelchair bound

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysphasic

ENVIRONMENTAL FACTORS

The patient is frequently driven in a car by her husband and volunteers. The car is used for holidays, stroke clubs and shopping trips. She needs to be able to give directions.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She is unable to give directions in the car.

COMMUNICATION ASSESSMENT - WITH AID

Although the patient was enthusiastic about the device she was unable to use it when the car was moving because she felt sick when reading the keys.

DEVICE RATING

The patient received a great deal of encouragement from her husband and from the Speech Therapist. The overall rating is low as the device could not be used in practice. She had no trouble operating the keys but felt sick looking at it when the car was moving.

OVERALL RATING:	*
EASE OF USE:	*
SOUND QUALITY:	*****
PORTABILITY:	*****
ENCOURAGEMENT:	*****
ACCEPTANCE:	**
VOCABULARY RANGE:	*****

NAME: M.H. **DEVICE:** Pocket Speech Aid (Personalised)

DATE OF BIRTH: 1931

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Dysphagia , wheelchair bound

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient is in long term hospital care.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient used gesturing and writing as her preferred mode of communication with other patients and with nursing staff.

COMMUNICATION ASSESSMENT - WITH AID

The patient did not use the device and preferred to gesture or write.

DEVICE RATING

The patient was not given enough encouagement in her routine environment. Nursing staff were able to anticipate her needs. Private conversation in the open ward was not possible using a speech output device.

OVERALL RATING: *
EASE OF USE: ***
SOUND QUALITY: ***
PORTABILITY: ***
ENCOURAGEMENT: *
ACCEPTANCE: *
VOCABULARY RANGE: *

NAME: R.W. **DEVICE:** Pocket Speech Aid (Personalised)

DATE OF BIRTH: 1931

MEDICAL DIAGNOSIS: Bilateral cerebrovascular accident

PHYSICAL DISABILITIES: Dysphagia , poor upper limb movement

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria, Good comprehension skills

ENVIRONMENTAL FACTORS

The patient is in long term hospital care. She is extremely sociable and has a great desire to communicate.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

She has poor finger movement but was able to use a Canon Communicator. The Canon did not allow her to attract and hold the attention of her carers.

COMMUNICATION ASSESSMENT - WITH AID

The aid improved her communication skills and gave her more control of her carers.

DEVICE RATING

The patient was enthusiastic about using the device but found it quite difficult to operate due to her limited finger movement. Portability was a problem as she was unable to carry the device. The vocabulary was personalised for her needs in long term hospital care. A larger number of phrases would be better for this patient and a single switch device easier to operate.

OVERALL RATING:	**
EASE OF USE:	***
SOUND QUALITY:	****
PORTABILITY:	*
ENCOURAGEMENT:	**
ACCEPTANCE:	***
VOCABULARY RANGE:	**

COMMUNICATION AID CASE STUDY

NAME:	J.R	**DEVICE:**	Pocket Speech Aid (peronalised)

DATE OF BIRTH: 1927

MEDICAL DIAGNOSIS: Pseudobulbar Palsy

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

This patient is a very active person and was hoping to return to work as a Draughtsman. He has frequent visitors to his home.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He uses a memowriter and writing to communicate because he requires an unlimited vocabulary. He is frustrated by his inability to control his young son.

COMMUNICATION ASSESSMENT - WITH AID

This patient used the Pocket Speech aid to discipline his son. This is something he found difficult to achieve without using speech. He was also issued with a Votrax to satisfy his need for a full range device.

DEVICE RATING

The device served the purpose it was intended for in this situation. He found the device easy to operate with acceptable sound quality. His listener also accepted the aid. Encouragement to use the device was provided by his wife.

OVERALL RATING:	****
EASE OF USE:	****
SOUND QUALITY:	****
PORTABILITY:	*****
ENCOURAGEMENT:	****
ACCEPTANCE:	***
VOCABULARY RANGE:	***

COMMUNICATION AID CASE STUDY

NAME: N.M. **DEVICE:** Mcleod unit

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

The patient lived at home with his family. He had a great desire to communicate. He accepted his prognosis and when he realised he was losing his speech he recorded several hundred words on cassette tape in the hope that these words could be transferred to a speech aid which he would be able to use when his voice deteriorated.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

As his disease was progressing his speech was becoming less intelligible but he still attempted to communicate using slow, slurred speech supplemented by gesturing.

COMMUNICATION ASSESSMENT - WITH AID

This patient was unusual in that he got a communication aid built to his own specifications. He designed the vocabulary and coding system for the device and was highly motivated to use the aid.

DEVICE RATING

The overall communication rating for the device is high. The device met all of this patient's specifications for a communication aid.

OVERALL RATING:	*****
EASE OF USE:	****
SOUND QUALITY:	**
PORTABILITY:	****
ENCOURAGEMENT:	*****
ACCEPTANCE:	*****
VOCABULARY RANGE:	****

NAME: D.R. **DEVICE:** Mcleod unit

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia, Limb weakness

MENTAL DISABILITIES: None

STATE OF SPEECH: Severe dysarthria

ENVIRONMENTAL FACTORS

The patient spent most of his time at home. He lived with his mildly handicapped son. He was a retired surgeon and was able to accept the prognosis.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He used writing and gesturing as his preferred mode of communication.

COMMUNICATION ASSESSMENT - WITH AID

He was frustrated by the slow communication rate and the limited vocabulary. The code chart was re-designed for faster access and the patient provided with his own custom vocabulary but he still used writing because it was quicker.

DEVICE RATING

The patient was enthusiastic about the speech aid but as it could not increase his communication rate it was rejected. The overall communication rating is low. The learning time of the code chart is too prolonged for a patient with a fast progressing disease.

OVERALL RATING: *
EASE OF USE: *
SOUND QUALITY: ****
PORTABILITY: ****
ENCOURAGEMENT: ***
ACCEPTANCE: *
VOCABULARY RANGE: **

NAME: M.M. **DEVICE:** Mcleod unit

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: None

ENVIRONMENTAL FACTORS

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient did not communicate at all. He made no effort to gesture or to write.

COMMUNICATION ASSESSMENT - WITH AID

The patient attempted to communicate using the device. The Mcleod unit provided strong motivation for the patient and gave him the appreciation of the use of speech for communication. The device provided the psychological factor neccessary for this patient who now has some functional language using gesturing and a few words.

DEVICE RATING

The overall rating in this case is high. The patient is now communicating whereas before introducing the aid he did not even attempt to communicate.

OVERALL RATING: ****
EASE OF USE: ***
SOUND QUALITY: ***
PORTABILITY: ***
ENCOURAGEMENT: ****
ACCEPTANCE: ****
VOCABULARY RANGE: ***

COMMUNICATION AID CASE STUDY

NAME: S.P. **DEVICE:** Mcleod unit

DATE OF BIRTH: 1926

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia, able-bodied

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient lives at home. Her previous occupation was as a clerkess/typist. She has good hand movement and is therefore able to operate keyboard devices efficiently. She leads a normal active life.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her preferred mode of communication is writing and gesturing. She occasionally uses the Votrax, particularly when communicating with her young grand-children who are unable to read.

COMMUNICATION ASSESSMENT - WITH AID

The patient prefers to write and gesture in normal conversation but finds the device useful for telephone communication.

DEVICE RATING

The patient likes the voice quality of the device which is superior to her previous device (Votrax). The coded input is too slow and leads to frustration. The patient requires a full range vocabulary for normal conversation and dislikes the limited vocabulary of the device. She uses the Mcleod unit as a telephone aid where the vocabulary restriction is not so critical.

	conversation	**telephone**
OVERALL RATING:	*	****
EASE OF USE:	*	***
SOUND QUALITY:	****	****
PORTABILITY:	*****	****
ENCOURAGEMENT:	***	****
ACCEPTANCE:	**	****
VOCABULARY RANGE:	**	****

COMMUNICATION AID CASE STUDY

NAME: S.P. **DEVICE:** Votrax

DATE OF BIRTH: 1926

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia, able-bodied

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient lives at home. Her previous occupation was as a clerkess/typist. She has good hand movement and is therefore able to operate keyboard devices efficiently. She leads a normal active life.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

Her preferred mode of communication is writing and gesturing.

COMMUNICATION ASSESSMENT - WITH AID

The patient prefers to write and gesture in normal conversation. Initially, she found the device to be a useful communication aid but she eventually grew tired of the mechanical sounding voice. After the first month she resorted to using the device for emergencies only. She occasionally uses the device to communicate with her grand-children as they enjoy hearing the robotic voice and are unable to read.

DEVICE RATING

Initially, the device was useful but is now used only occasionally. She did not like the mechanical sounding voice. Her disease has now progressed and she is less able to operate the keyboard. She is still able to operate the simpler Pocket Speech Aid device.

OVERALL RATING:	**
EASE OF USE:	**
SOUND QUALITY:	*
PORTABILITY:	*
ENCOURAGEMENT:	*
ACCEPTANCE:	**
VOCABULARY RANGE:	***

COMMUNICATION AID CASE STUDY

NAME: J.R. **DEVICE:** Votrax

DATE OF BIRTH: 1927

MEDICAL DIAGNOSIS: Pseudobulbar Palsy

PHYSICAL DISABILITIES: Dysphagia

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

The patient is a very active person and was hoping to retun to his occupation of a Draughtsman. He has frequent visitors to his home.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He uses a memowriter and writing to communicate as he requires an unlimited vocabulary. He is frustrated by his inability to control his young son.

COMMUNICATION ASSESSMENT - WITH AID

The Votrax was introduced as a full range conversation aid for communicating at home. The device was rarely used due to the patients embarrasment of the synthetic voice. He preferred to communicate through writing as it was easier and much quicker.

DEVICE RATING

The overall rating of the device is low because the patient did not like the poor sound quality and found the device too slow for conversation. Initially he was enthusiastic about the unlimited vocabulary provided by the Votrax but other factors had a greater influence on him.

OVERALL RATING: *
EASE OF USE: **
SOUND QUALITY: *
PORTABILITY: **
ENCOURAGEMENT: ***
ACCEPTANCE: **
VOCABULARY RANGE: *****

COMMUNICATION AID CASE STUDY

NAME: J.D. **DEVICE:** Votrax

DATE OF BIRTH: 1943

MEDICAL DIAGNOSIS: Multiple sclerosis

PHYSICAL DISABILITIES: Wheelchair bound

MENTAL DISABILITIES: None

STATE OF SPEECH: Dysarthria

ENVIRONMENTAL FACTORS

The patient spends all of his time at home and depends on his family to meet his needs.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He used gesturing and dysarthric speech as his preferred mode of communication.

COMMUNICATION ASSESSMENT - WITH AID

The patient rejected the device immediately. He did not like the mechanical sounding voice and it was difficult for him to operate due to his limited finger movement.

DEVICE RATING

The device was unsuitable for this patient.

OVERALL RATING: *
EASE OF USE: *
SOUND QUALITY: *
PORTABILITY: *
ENCOURAGEMENT: *
ACCEPTANCE: *
VOCABULARY RANGE: ***

COMMUNICATION AID CASE STUDY

NAME: M.H. **DEVICE:** Votrax

DATE OF BIRTH: 1931

MEDICAL DIAGNOSIS: Cerebrovascular accident

PHYSICAL DISABILITIES: Dysphagia, wheelchair bound

MENTAL DISABILITIES: None

STATE OF SPEECH: Anarthria

ENVIRONMENTAL FACTORS

The patient is in long term hospital care.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient used gesturing and writing as her preferred mode of communication with other patients and with nursing staff.

COMMUNICATION ASSESSMENT - WITH AID

The patient used the device for approximately one month. She used the device to communicate with patients and nursing staff. She eventually grew tired of the device and rejected it.

DEVICE RATING

The patient was not given enough encouagement in her routine environment. Nursing staff were able to anticipate her needs. Private conversation in the ward was not possible. The device was a useful communication aid for the first month.

OVERALL RATING: **
EASE OF USE: **
SOUND QUALITY: *
PORTABILITY: **
ENCOURAGEMENT: *
ACCEPTANCE: *
VOCABULARY RANGE: ***

NAME: D.R. **DEVICE:** Votrax

DATE OF BIRTH:

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: Dysphagia, Limb weakness

MENTAL DISABILITIES: None

STATE OF SPEECH: Severe dysarthria

ENVIRONMENTAL FACTORS

The patient spent most of his time at home. He lived with his mildly handicapped son. He was a retired surgeon and was able to accept the prognosis.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

He used writing and gesturing as his preferred mode of communication.

COMMUNICATION ASSESSMENT - WITH AID

The patient rejected this device almost immediately. He did not like the mechanical sounding voice.

DEVICE RATING

The overall communication rating is low. The patient did not like the artificial voice and he found the device too slow for adequate communication.

OVERALL RATING:	*
EASE OF USE:	**
SOUND QUALITY:	*
PORTABILITY:	**
ENCOURAGEMENT:	***
ACCEPTANCE:	*
VOCABULARY RANGE:	***

NAME: W.M. **DEVICE:** Uvocom

DATE OF BIRTH: 21.9.25

MEDICAL DIAGNOSIS: Motor Neurone Disease

PHYSICAL DISABILITIES: None

MENTAL DISABILITIES: None

STATE OF SPEECH: Unintelligible

ENVIRONMENTAL FACTORS

The patient leads an active life and requires to communicate in many different situations.

COMMUNICATION ASSESSMENT - PRIOR TO INTRODUCTION OF AID

The patient is an accomplished user of augmentative communication systems. She uses a word processing package on a home computer to generate messages for her family and friends and makes use of a memowriter when outside her home. She has a British Telecom 'Claudius Converse' for use on the telephone.

COMMUNICATION ASSESSMENT - WITH AID

She found the device to be of most benefit as a telephone aid. The unlimited vocabulary gave her more control of a conversation than the 'Cladius Converse'. She preferred to use the device in the interactive mode as it kept listeners interested.

DEVICE RATING

The patient is used to augmentative communication and has developed her own communication strategies over a long period of time. The device was a useful supplement to her collection of devices but she did not feel a need for speech in preference to the written word in most situations.

OVERALL RATING: **
EASE OF USE: ***
SOUND QUALITY: ***
PORTABILITY: *****
ENCOURAGEMENT: *****
ACCEPTANCE: *****
VOCABULARY RANGE: *****

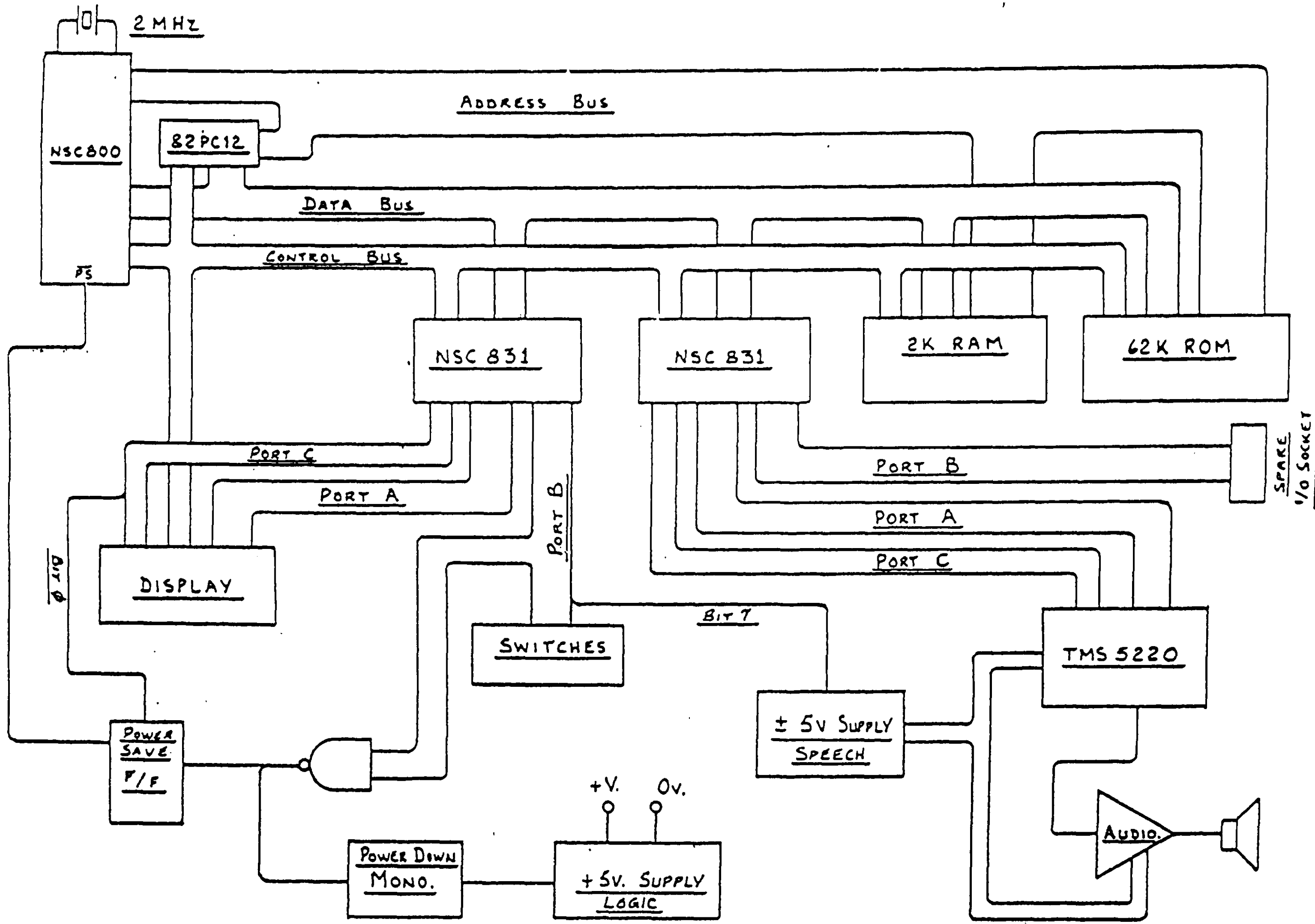

FIGURE A7.1 The Macleod Unit: Schematic Diagram

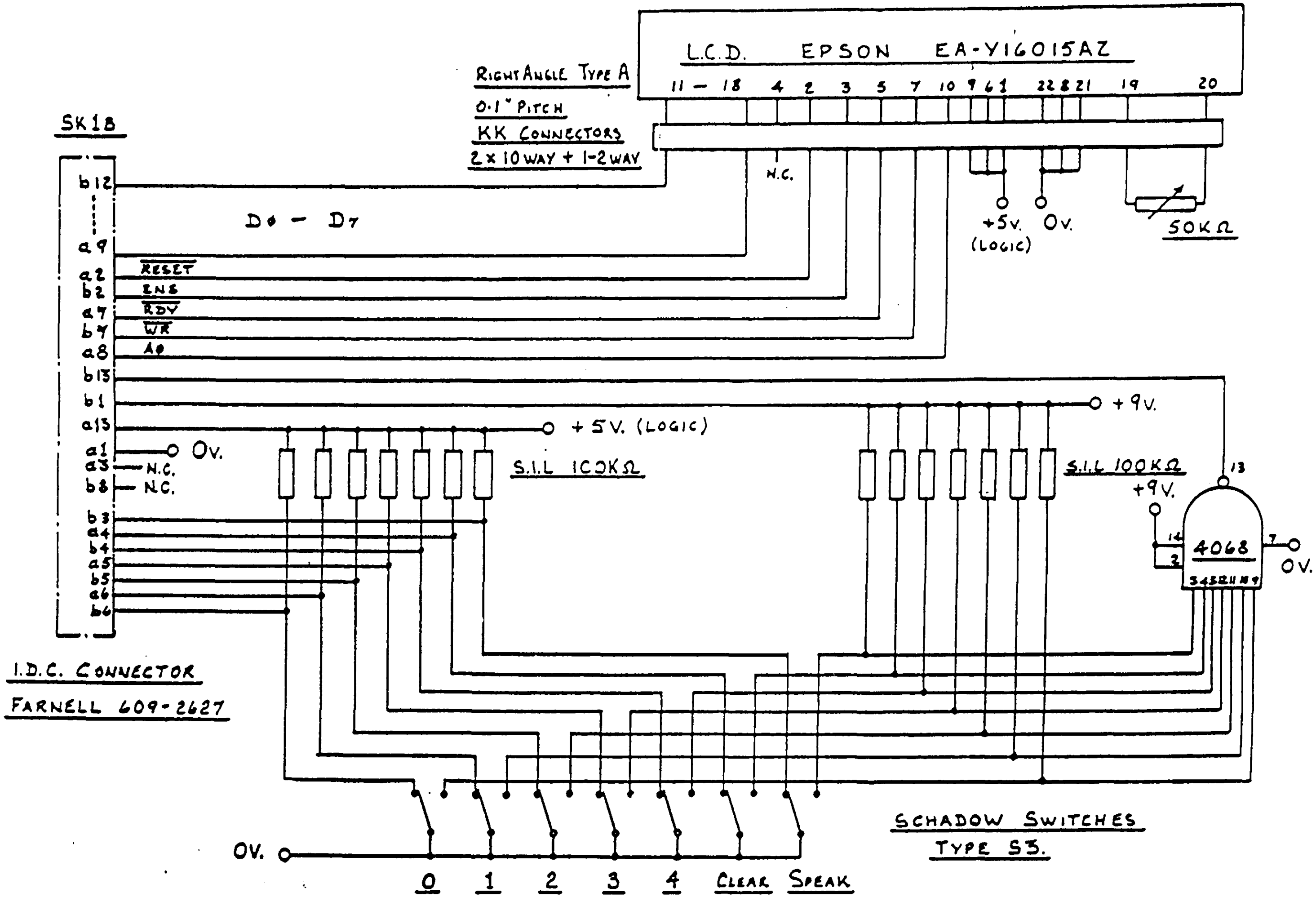

FIGURE A7.2 The Macleod Unit: Switch Input Circuit

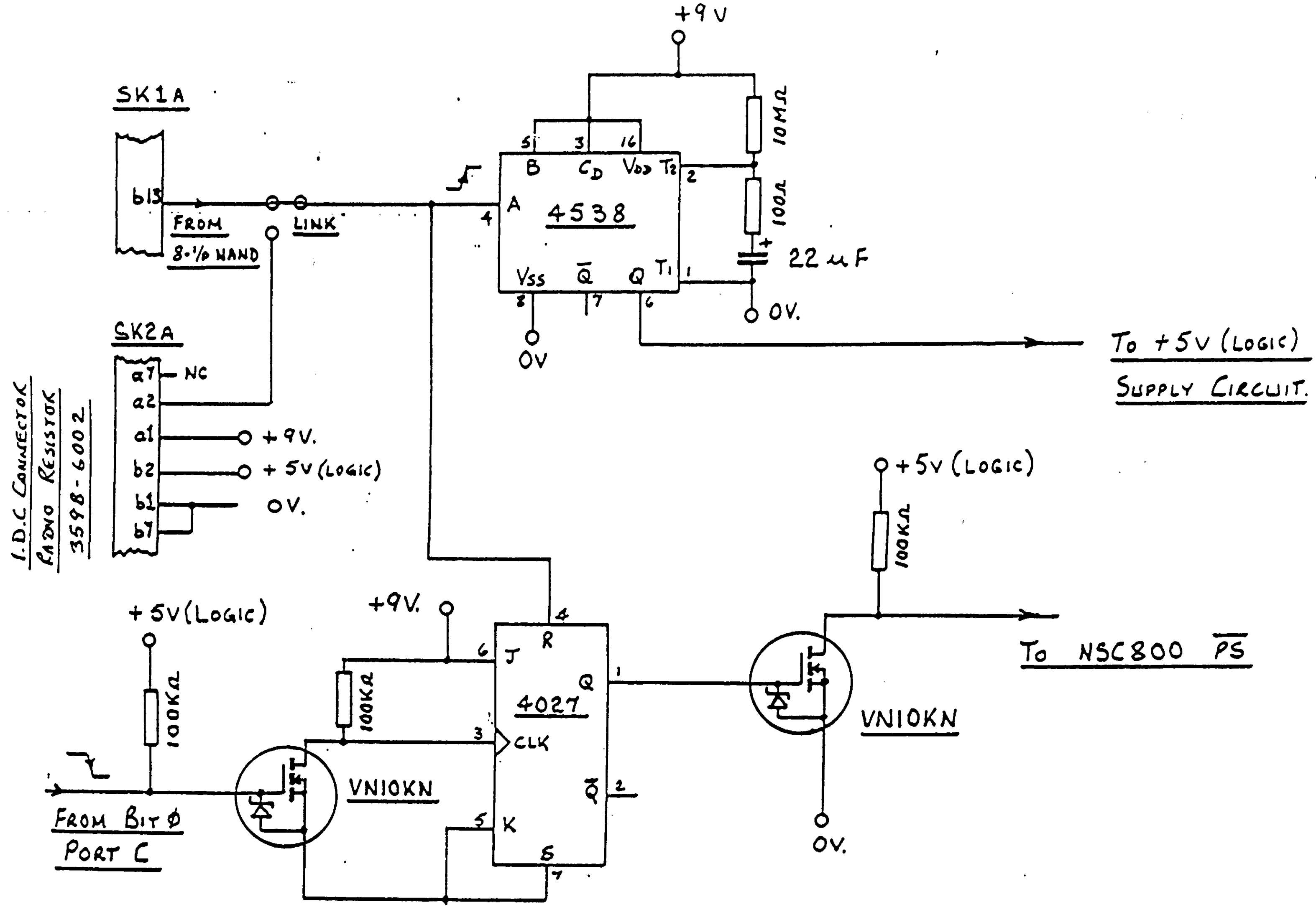

FIGURE A7.3 The Macleod Unit: Power Up and Power Save Circuit

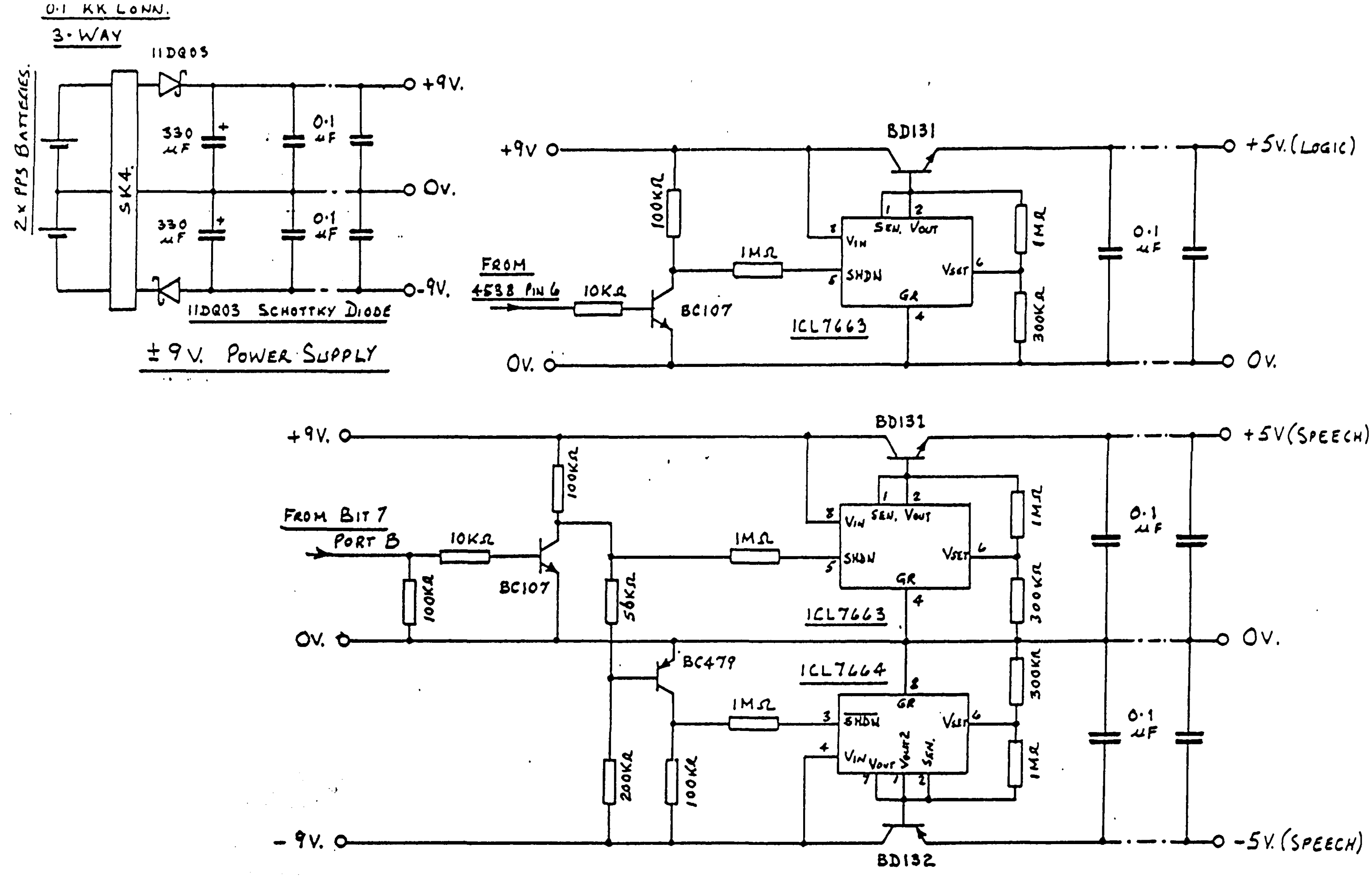

FIGURE A7.4 The Macleod Unit: Power Supply Circuits

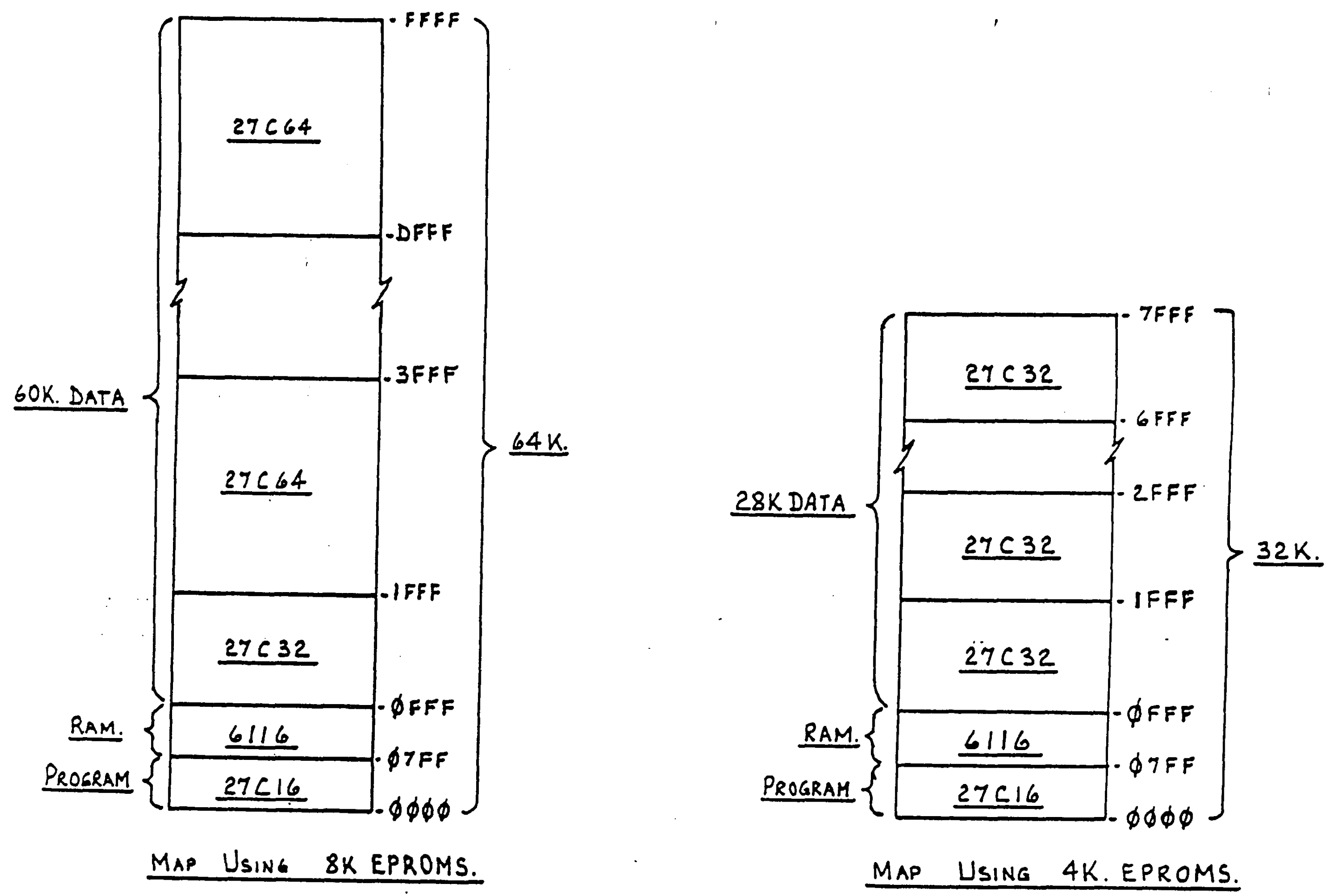

FIGURE A7.5 **The Macleod Unit: Memory Map**

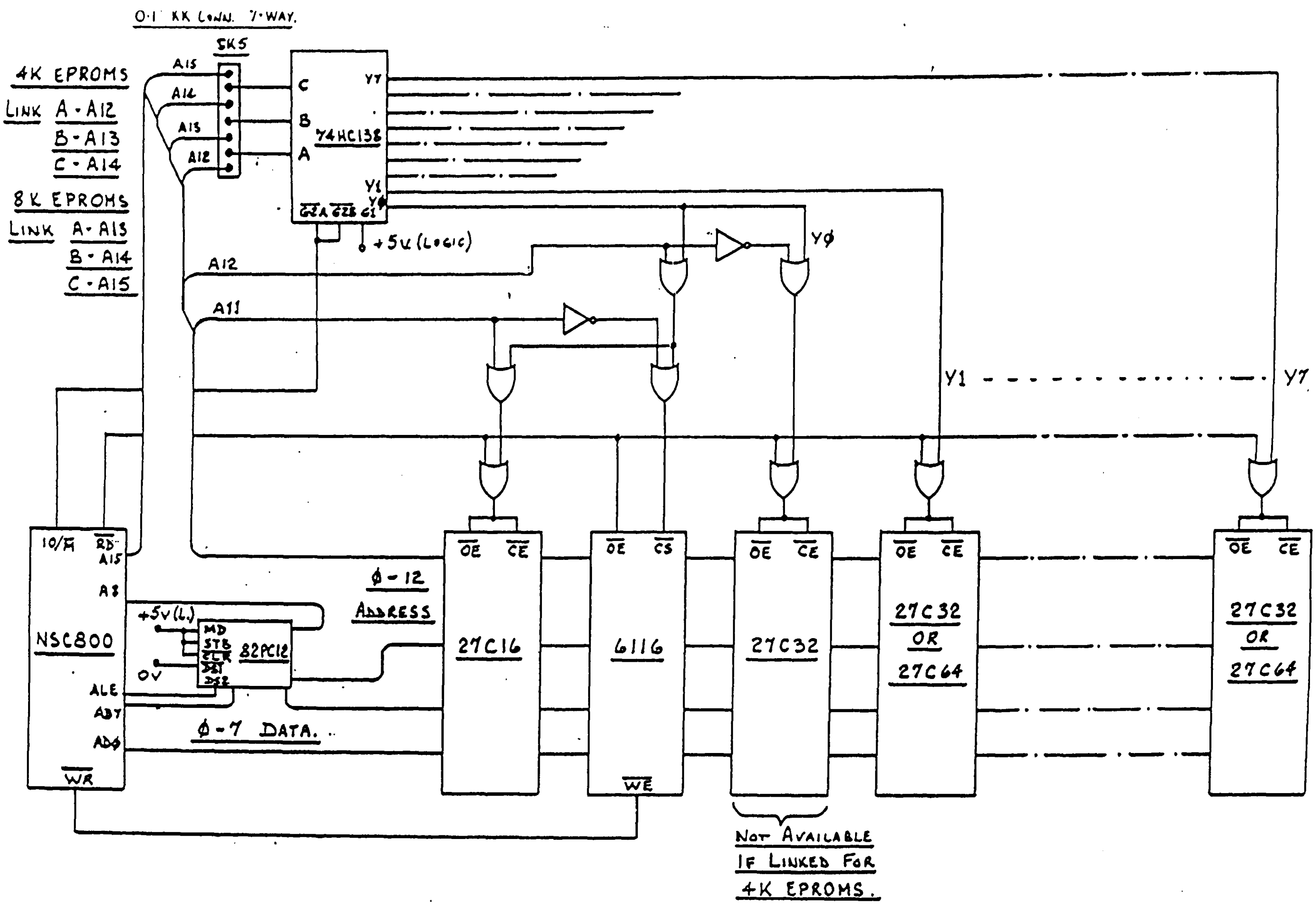

FIGURE A7.6 The Macleod Unit: Address Decoding Circuit

APPENDIX 7 - THE MACLEOD UNIT

A schematic diagram of the Macleod unit is shown in Figure A7.1 and a flow-chart of the main programme loop is shown in Figure A7.10.

Start Up

A circuit diagram of the switch input circuit is shown in Figure A7.2.

Pressing any key will take an input of the 8 input NAND gate (4068) low and result in a low to high transition on its output. The rising edge will trigger the monostable (4538) to output a 2 minute duration retriggerable pulse. The output from the 4538 is taken via a BC107 transistor inverter to the shut down control of the ICL7663 voltage regulator (Figure A7.3). The low signal switches on the regulator and gives via the BD131 series transistor the +5V logic supply. The power supply circuits are shown in Figure A7.4. Turning ON the +5V supply will power up all memory, I/O ports, display, decoding logic and the NSC800 microprocessor. A 2 MHz crystal provides the clock for the processor and a 1 MHz output from the NSC800 is used by the display. An RC network resets the processor and starts the program at location 0000 running.

Memory Addressing

A memory map is shown in Figure A7.5.

With the system set up to the full 64K i.e links A15-C, A14-B AND A13-A on and SK5 in place, address decoding is as follows (Figure A7.6). The low eight address lines are demultiplexed from the AD outputs of the NSC800 by the 82PC12. These lines together with

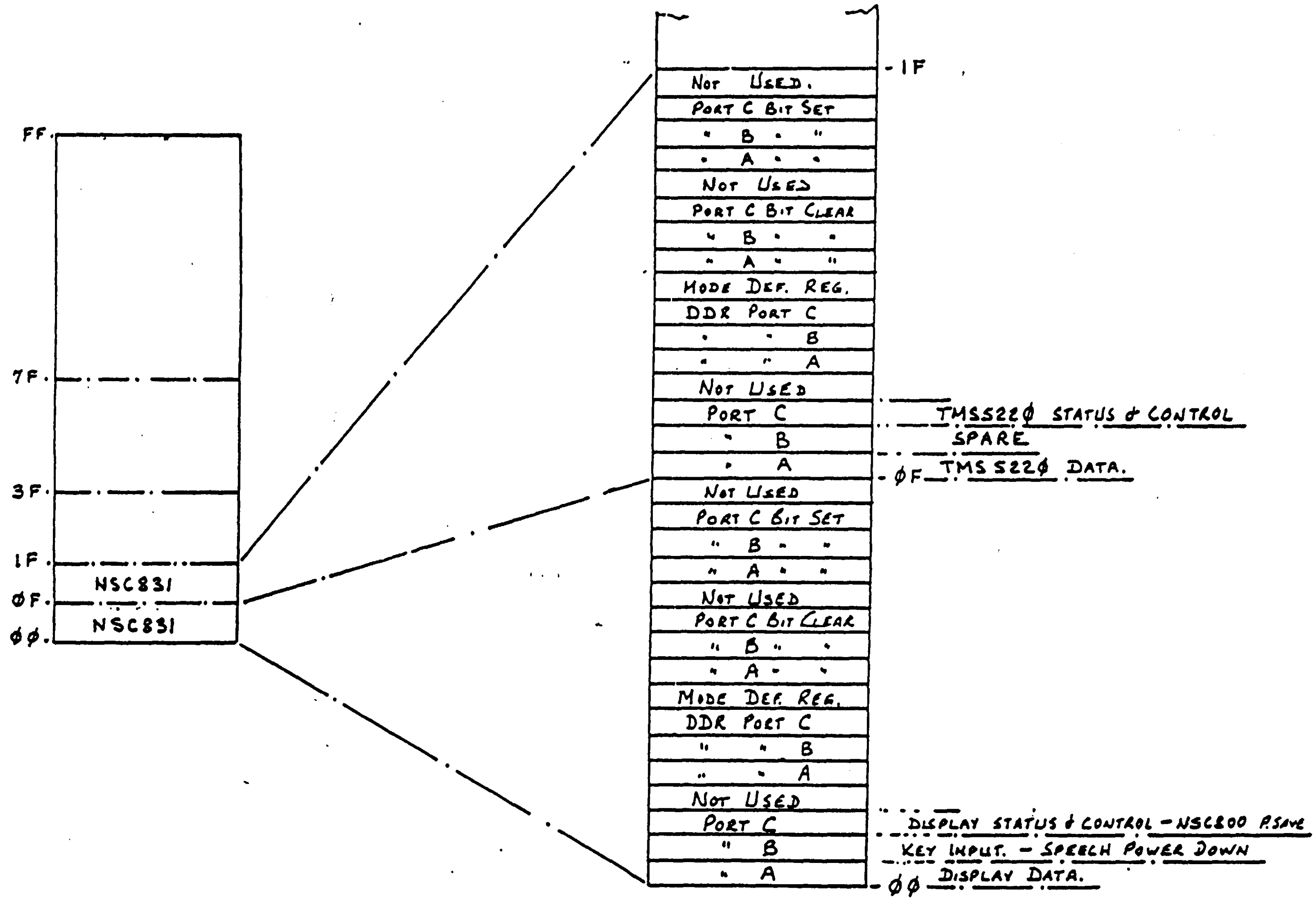

FIGURE A7.7 The Macleod Unit: Input/Output Memory Map

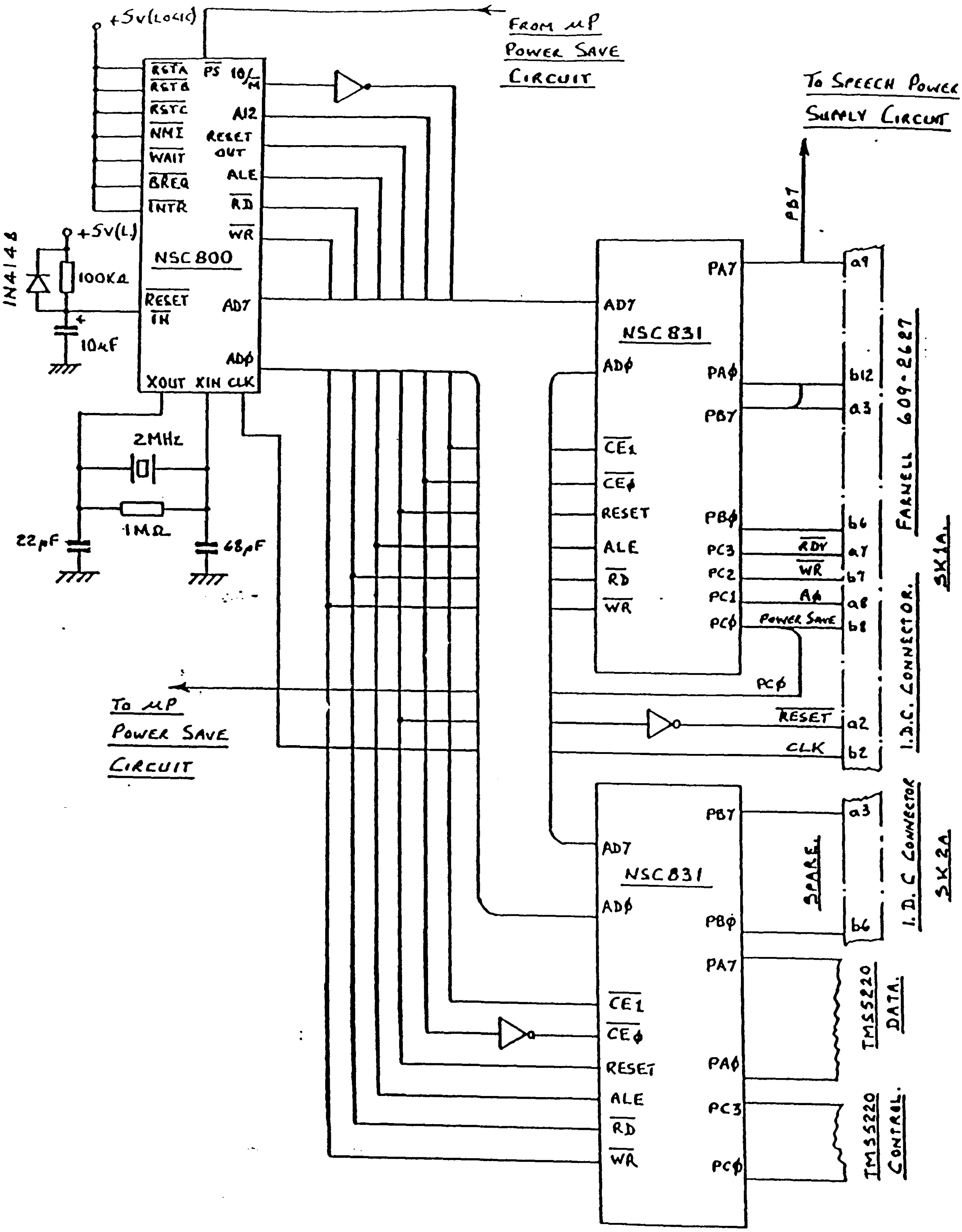

FIGURE A7.8 The Macleod Unit: Input/Output Ports the upper 8 address lines form the 16 lines A0-A15 that are used in memory addressing. Lines A13-A15 and line IOM, the memory select line, will decode using the 3 to 8 decoder 74HC138 for an 8K slot in memory. For locations above 1FFF this will be a 27C64 EPROM. Further decoding and chip select is by lines A0-A12 and the read line RD. For locations 0000 to 1FFF the 8K slot will be occupied by three memory devices, a 6116 RAM, a 27C16 EPROM and a 27C32 EPROM. Selection of the device is by lines A11 and A12 and within the device by lines A0-A11. A further line, the write enable WR, is taken to the 6116 RAM.

I/O Addressing

The I/O memory map is shown in Figure A7.7 and the circuit for the I/O ports is shown in Figure A7.8.

With the program running, the two NSC831 devices are configured for input or output on the various ports. Port and bit allocation is as follows:

Port Address 00	-	Bits 0-7 output to Display.
Port Address 01	-	Bits 0-6 input from Key Switches.
	-	Bit 7 output to control +5V Speech Supply.
Port Address 02	-	Bit 0 output NSC800 power save control
	-	Bits 1 and 2 output display status and control
	-	Bit 3 input display status
Port Address 10	-	Bits 0-7 input and output for TMS 5220
Port Address 11	-	Bits 0-7 spare port
Port Address 12	-	Bits 0 and 1 input TMS 5220 status and control

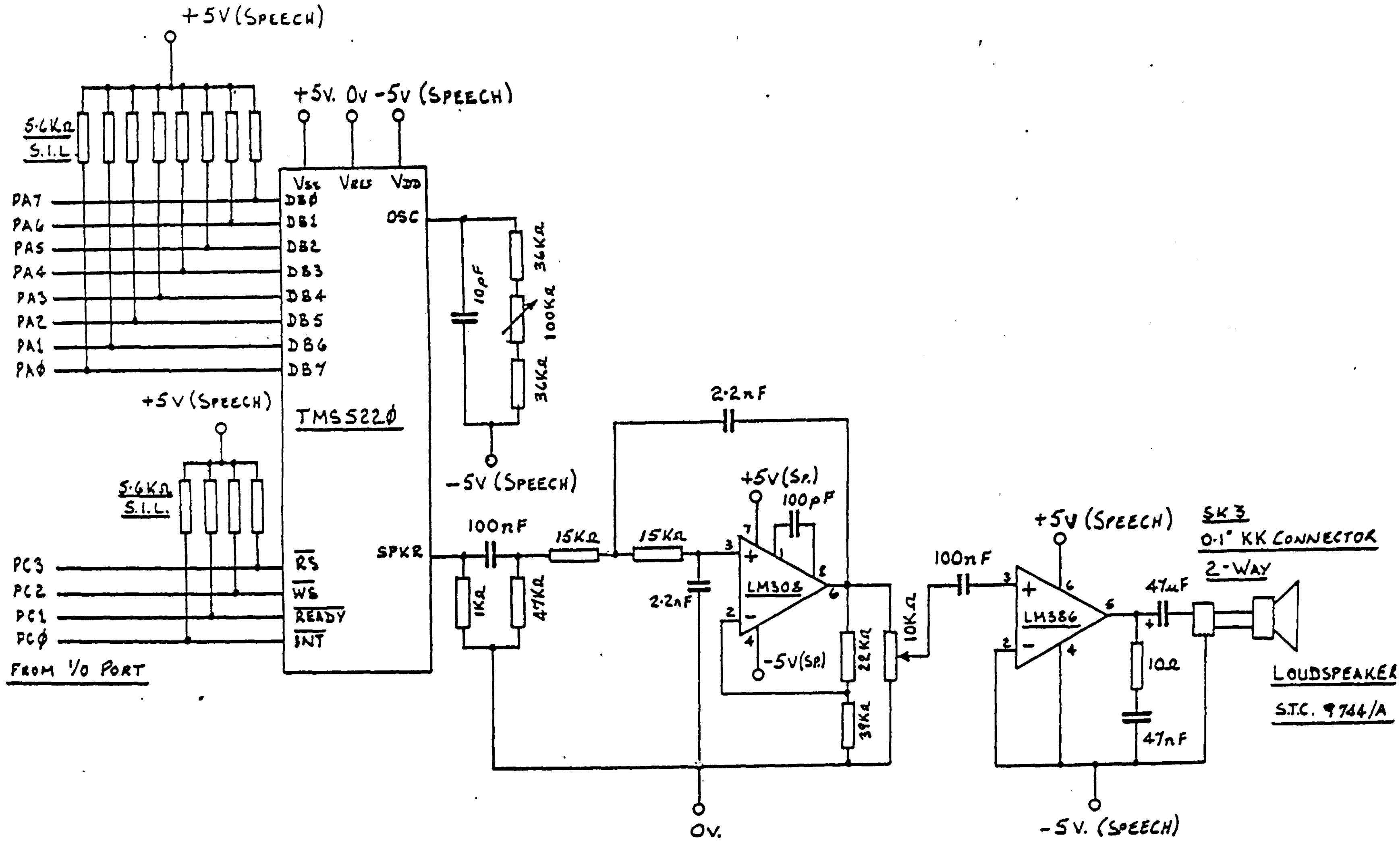

FIGURE A7.9 The Macleod Unit: Processor and Audio Circuit

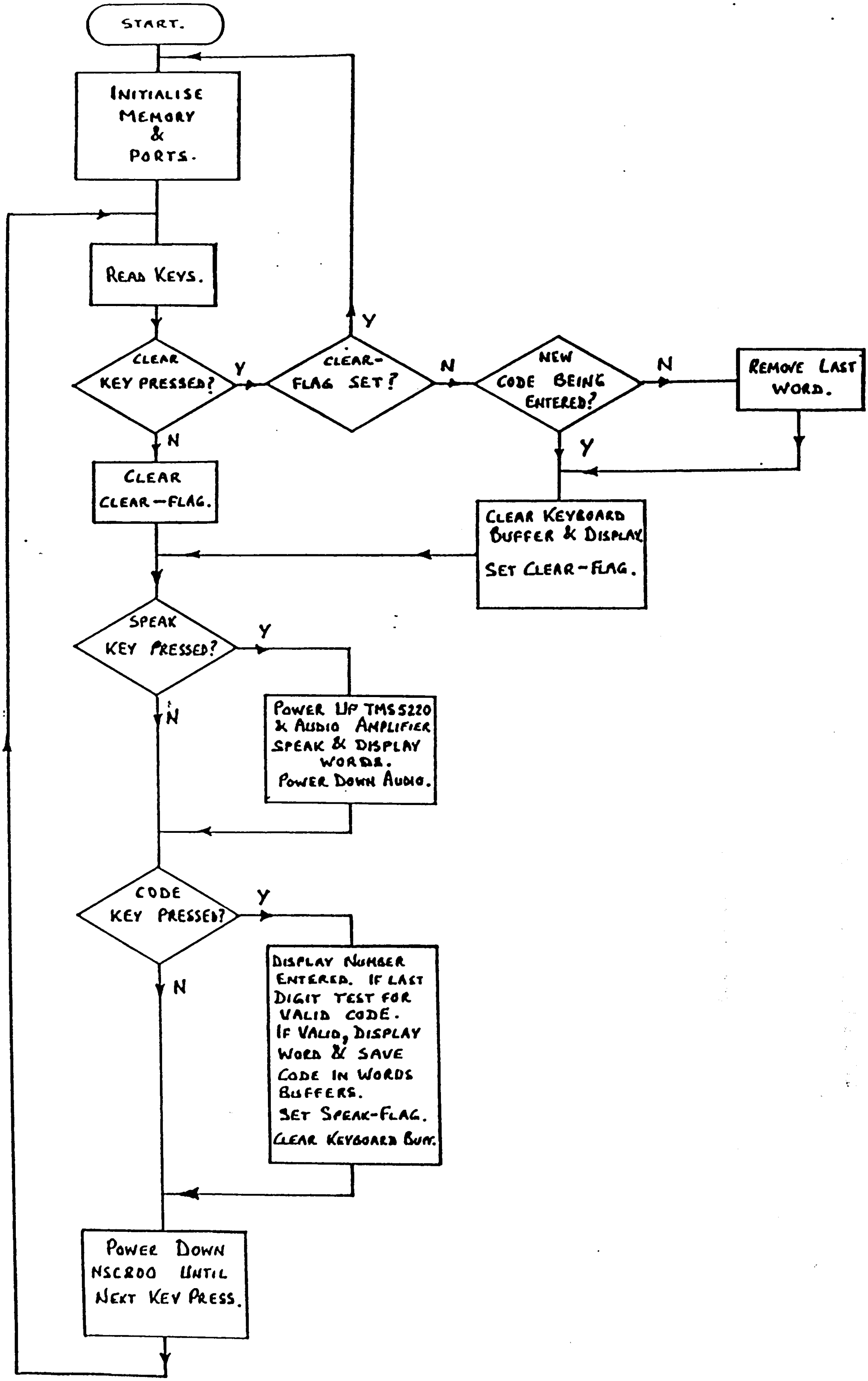

FIGURE A7.10 The Macleod Unit: Main Program Flow Chart

- Bits 2 and 3 input and output TMS 5220 status and control

The program will now use port 00 bits 0-7 and port 02 bits 1-3 to ready the display. Finally the program enters the keyboard read routine.

Key Read

A flow-chart of the code entry and verification programme is shown in Figure A7.11. The programme sits waiting for a code key to be pressed. When this is done the first digit of the 4 digit code is displayed and the entered code saved. The program will now drop through to a power save routine. This routine will pulse bit 0, port 02, low. A low on bit 0 will turn off VFET VNIOKN and the resulting low to high transition will clock a logic high to the Q output of the 4027 latch (Figure A7.3). The latch output will turn ON the second VFET causing a logic low on the NSC800 PS input. On completion of its current instruction the processor will enter a power save mode.

Entering the next code digit will cause the following sequence of events. The 4068 will go high triggering the 4538 monostable and thus keeping the +5V logic supply ON. It will also reset the 4027 latch. The resulting low on the latch output will turn OFF the VFET and remove the low on the NSC800 PS input. The processor now continues with the program, which is a jump to the key input port. The entered code will be displayed and saved and the above cycle repeated. If the digit entered is the fourth digit of the code then before going into the power save routine the program will carry out a code verification and conversion routine. If the code is

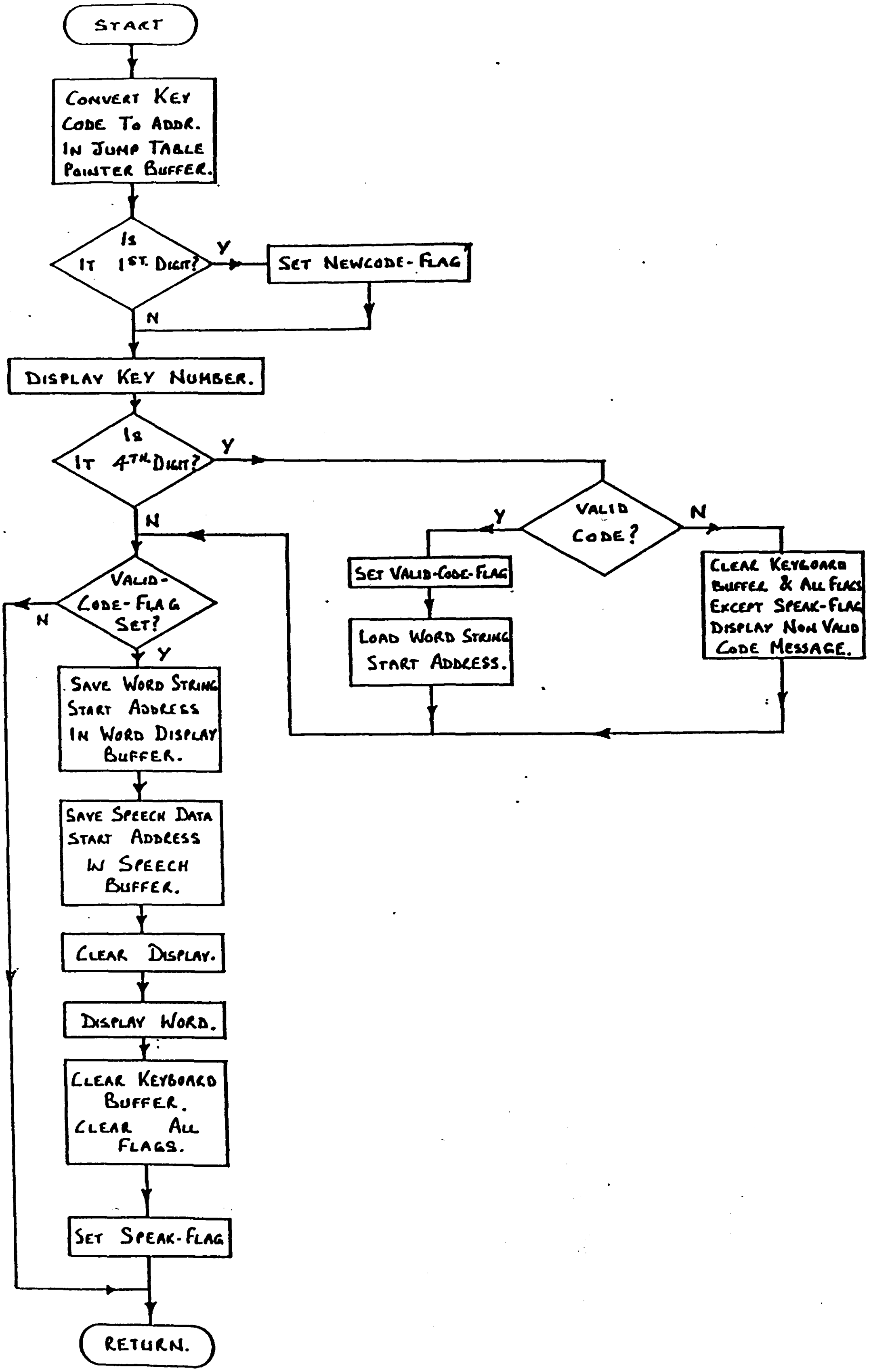

FIGURE A7.11 The Macleod Unit: 'Code Entry and Verification' Flow Chart

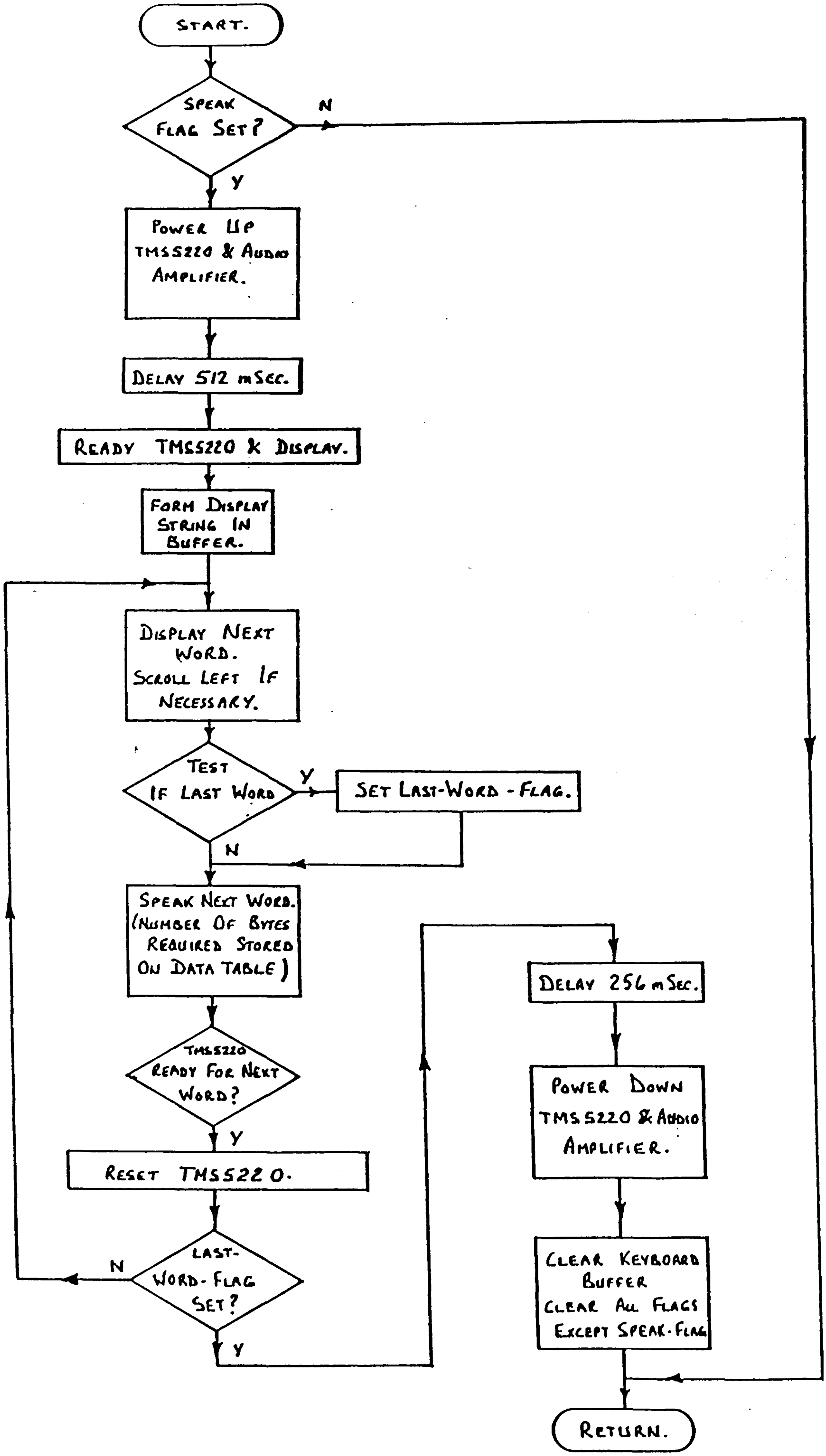

FIGURE A7.12 The Macleod Unit: ´Speak and Display Words` Flow Chart valid then the corresponding word for this code will be displayed and the code coverted to an entry in a data jump table. The entry address is then saved in the words buffer. Should the entered code be non-valid then a 'Not valid code' message will be displayed and all keyboard entries cleared.

If at any time the speak key is pressed the program and circuit will carry out the following actions. As with the code keys the +5V logic supply will be retriggered and the NSC800 will come out of the power save mode. On reading the keys port and finding that the speak key was pressed the program enters a speak routine. First it tests to see if words have been entered. If this is not the case then the program drops through into the power save routine.

However, if data has been entered then the audio circuit is powered up. A high is output on bit 7 of port 01. This will turn ON voltage regulators ICL 7663 and ICl 7664 via transistors BC107 and BD132. By turning on the regulators the +5V speech supply is obtained through transistors BD131 and BD132. There is now a delay of 500 msec to allow the power supplies to stabilise.

Speak and Display

The processor and audio cicuit is shown in Figure A7.9 and a flow-chart of the programme is shown in Figure A7.12.

After the audio power up stabilising delay the program will fetch and display the text for the first word.

The TMS 5220 is then sent the reset code putting it into a known state. This is followed by the speak external code in order to allow data to be supplied to the TMS 5220 over its data bus. The program will now fetch the first two bytes of data and interpreting this as a 16 bit binary number will proceed to pass that number of speech data bytes to the speech processor. After the last data byte has been passed the program will test for any more words. If there are more words then the next word will be displayed, scrolling the exising word left if neccessary. The TMS 5220 chip is then tested for completion of speaking the current word. On completion it will be reset, set for speak external and the data for the next word sent as before. If there are no more words then a 250 msec delay will allow completion of speaking the last word before power down of the audio circuit takes place. To power down a low is output on bit 7 of port 01. This signal via transistors BC107 and BC479 will shut down voltage regulators ICL 7663 and ICL 7664 thus removing the +5V speech supply. The speech output from the TMS 5220 processor is passed through a high pass passive filter to remove any DC signal component. The signal is then passed through a low pass second order active filter with a 5KHz cut-off.

The output from the LM308 op-amp of the filter is passed via a volume control (internal preset) and AC coupling to the LM386 audio amplifier and thus to the internally mounted loudspeaker.

Clear Key

From a hardware point of view, pressing the Clear Key will have the same effect as pressing any other key i.e. the 5V logic will be retriggered and the NSC800 will exit power save mode. The action taken will however depend on keys that were pressed previously.

On entering the clear routine the program checks if the last key pressed was the clear key and if so then all entered data is cleared and the system re-initialises as for start-up. If the last key pressed was not clear then the program tests for a new code being entered i.e. fourth digit not yet entered. If this is the case then the keyboard buffer is cleared ready for the new code entry. In all other cases pressing the clear key will remove the last entered word from the words buffer. After completion of one of the clear actions the program drops down into the power save routine and the NSC 800 is put in its power save mode.

Power Down

If no further key presses occur within a time interval of 2 minutes the system will switch itself OFF as follows. The 4538 monostable will time out and its output will go low. This signal, via the BC107 inverter, will cause the ICL 7663 regulator to shut down thus switching off the +5V supply. All memory, ports, processor etc will power down.

Only the key sensing logic, which is on the direct +9V supply from the battery, will remain powered up. The quiescent current required to run these CMOS devices is in the order of 80 uA and there is therefore a negligible drain on the battery.

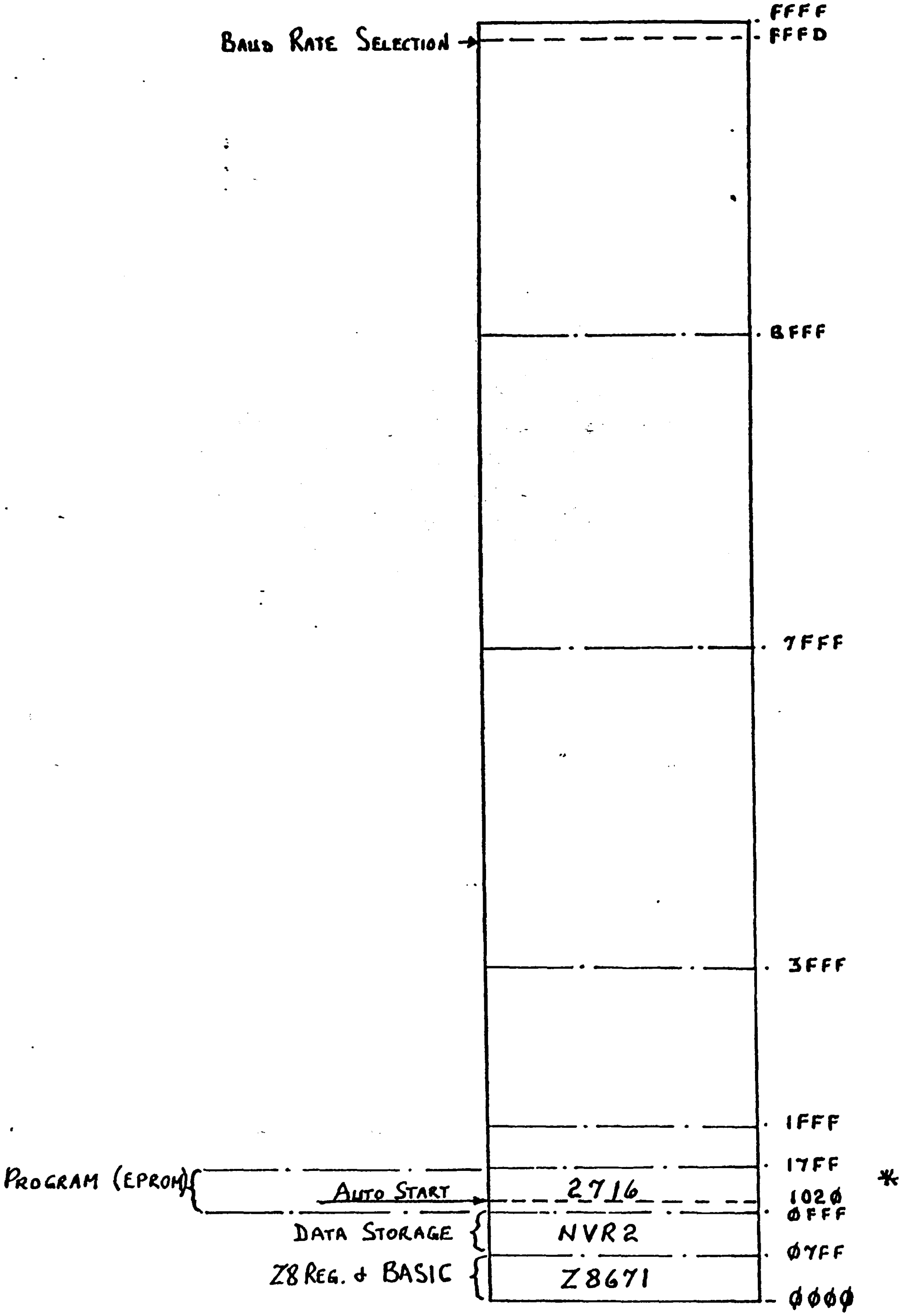

FIGURE A8.1 Text to Speech Mk I: Interface Memory Map

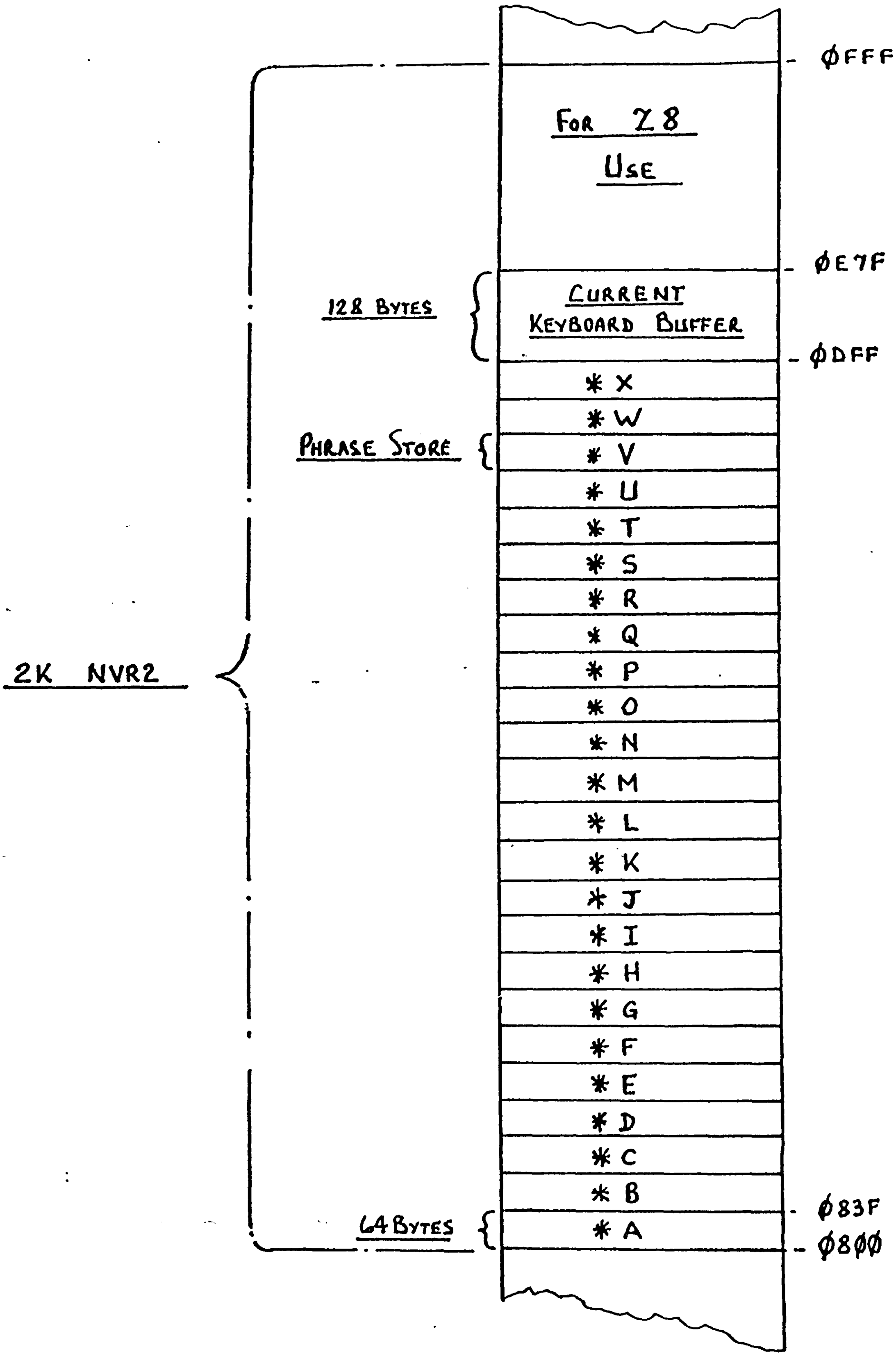

FIGURE A8.2 Text to Speech Mk I: NVR Memory Map

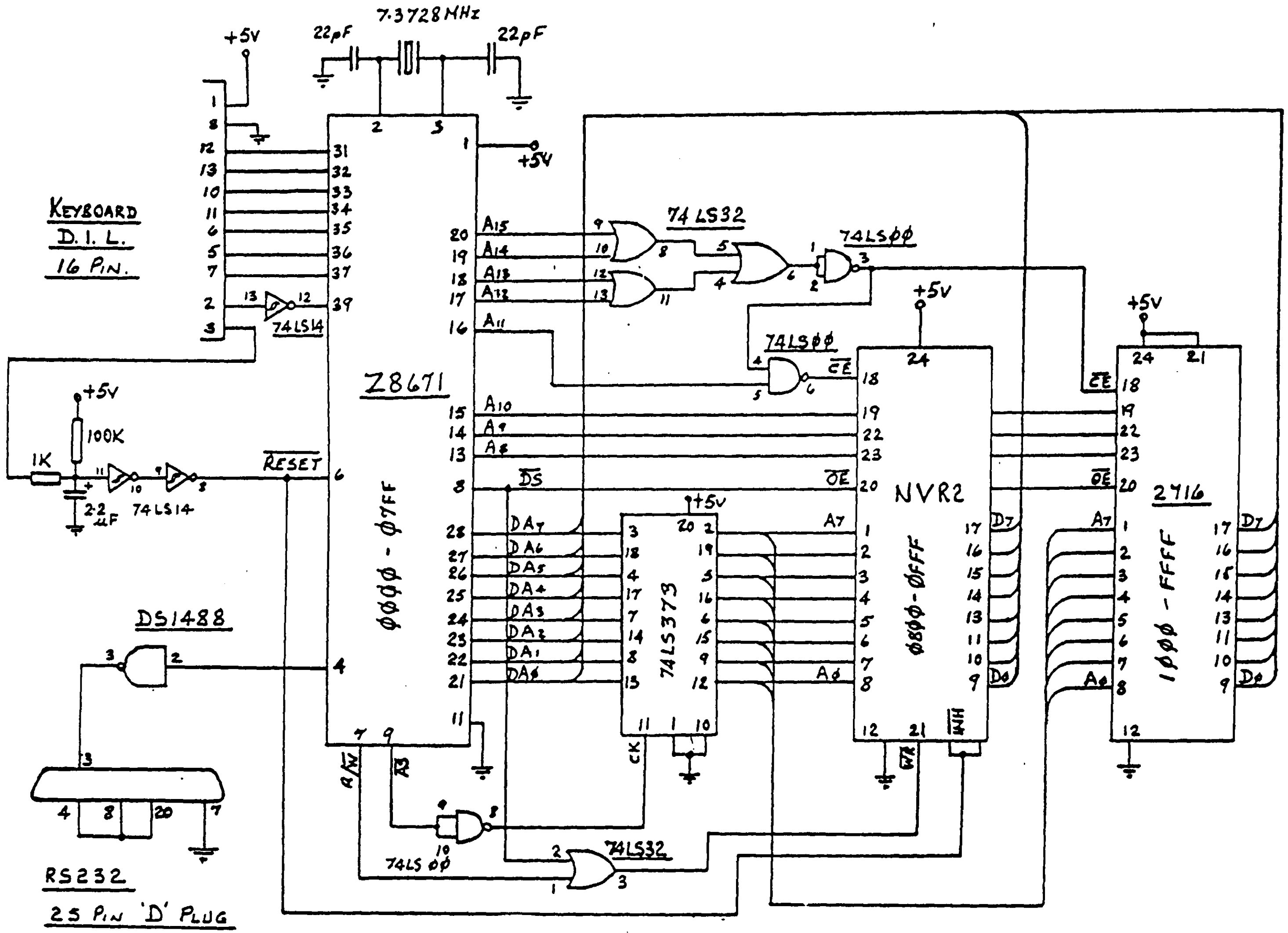

FIGURE A8.3 Votrax/Keyboard Buffer Interface

APPENDIX 8 - TEXT-TO-SPEECH Mk I

The microprocessor chip used is the Z8671 version of the Z8 processor and runs at 7.3728 MHz. It has 2K non-volatile RAM at locations 0800 to OFFF and a 2K EPROM at locations 1000 to 17FF. The EPROM has wrap-around addressing and is reflected all the way up to FFFF. The Z8671 uses locations 0000 - 07FF for its own use. The memory map of the interface is shown in A8.1 and the NVR memory map is shown in A8.2. The Votrax keyboard buffer interface is shown in A8.3. On power up a POC signal resets the Z8671 and takes the inhibit input of the NVR low, this prevents writing to the NVR until all system signals are established. The processor then searches for the existence of RAM and notes its location in the memory map. The Z8671 requires part of the RAM for its own use.

The processor then examines FFD (this is in fact location 17FD) in order to set the required baud rate. Finally it in inspects location 1020 for the existence of an EPROM stored program and auto-starts on finding it at this location.

Pressing any key will generate a logic high strobe signal. This signal is inverted to give the DAV signal for Port 2 in the Z8671. The DAV signal latches the data on the databus into Port 2. When this is done the status bit for Port 2, i.e. bit D2, in the Z8671 register No. R250 is set. The program polls this register and on finding it set transfers the data in Port 2 into memory. It then clears the status but ready for the next key entry.

The program now examines the entered character and depending on the ASCII character may take one of several actions.

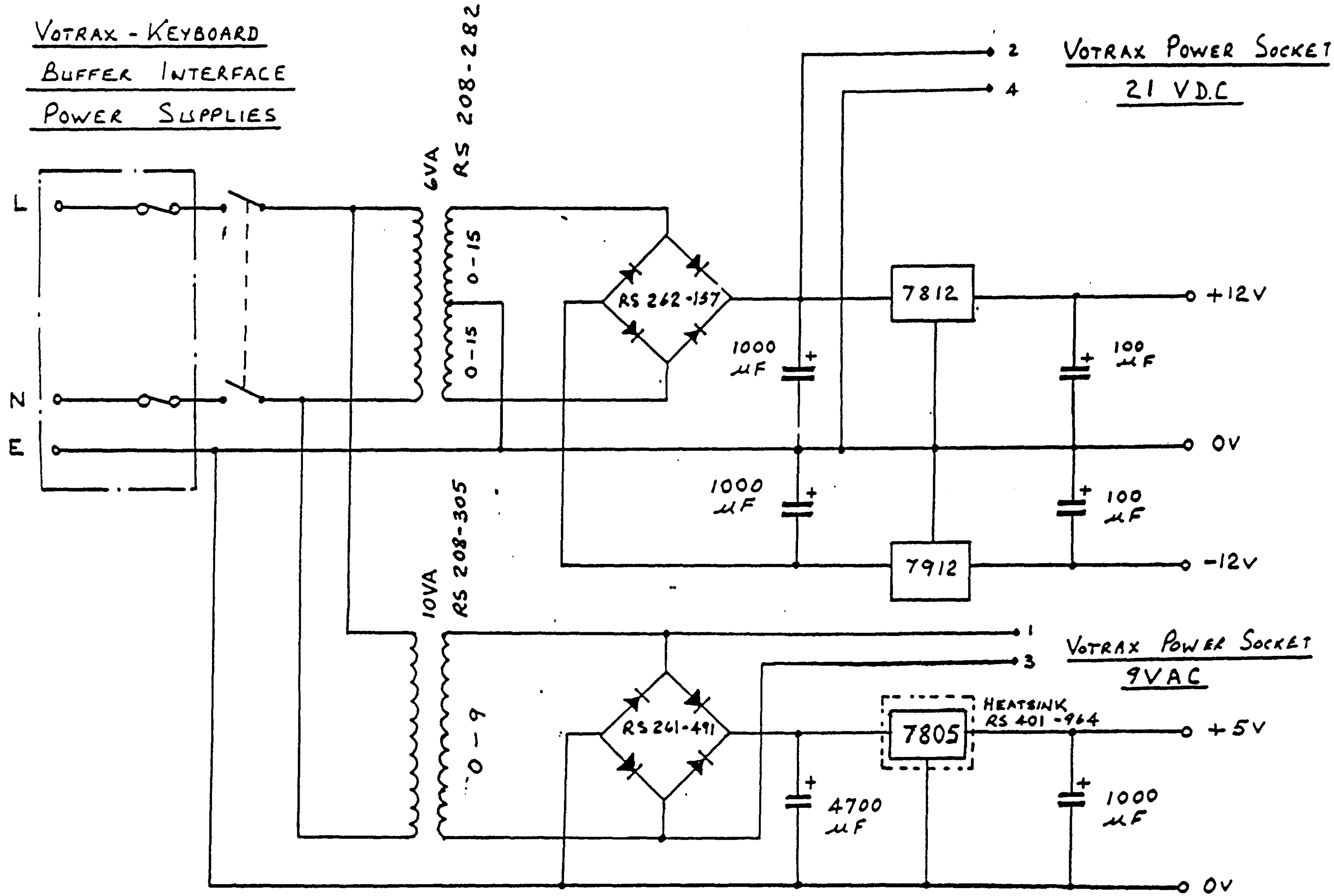

FIGURE A8.4   Text to Speech Mk I: Power Supply Circuit

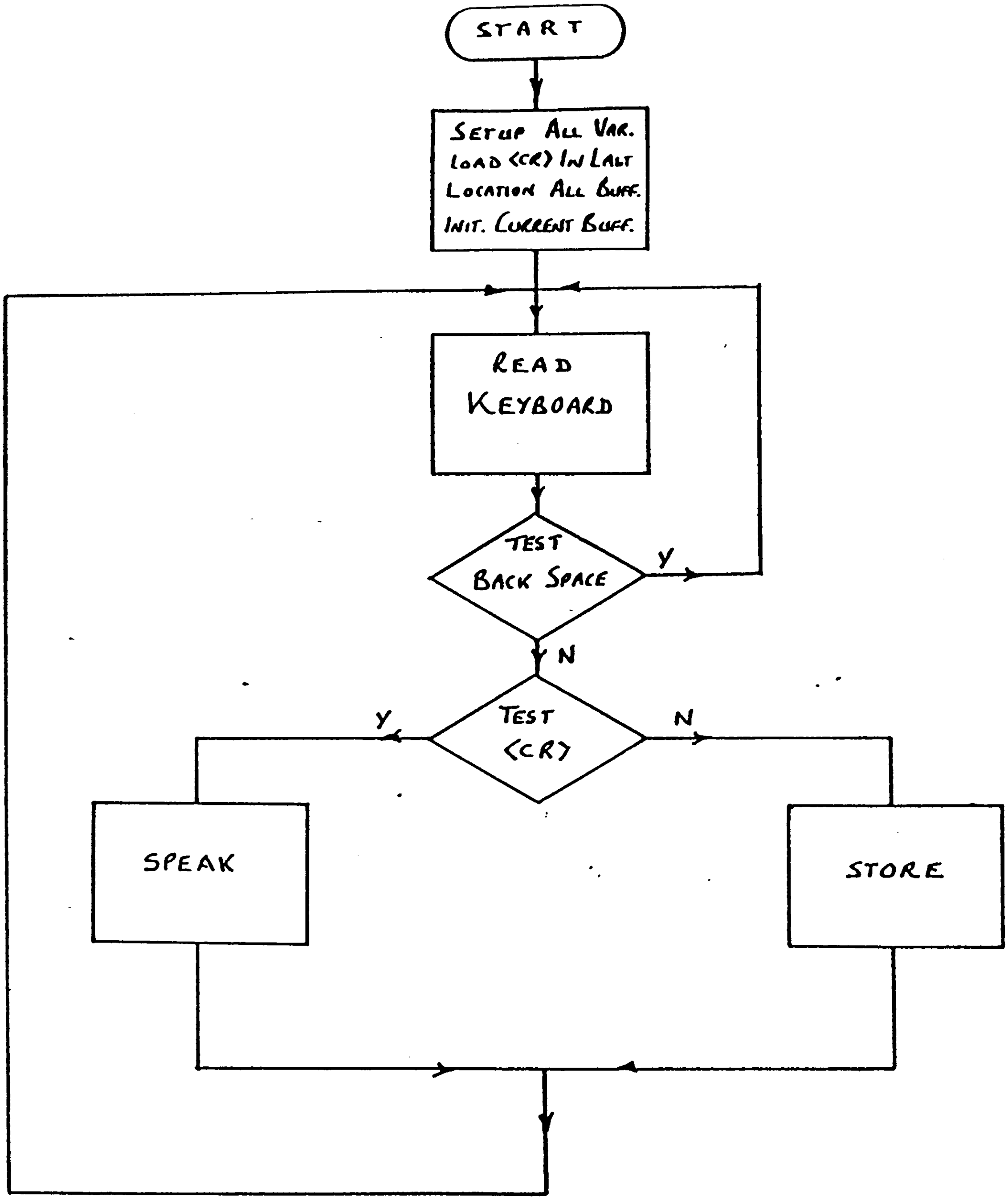

FIGURE A8.5 Text-to-Speech Mk I: Main Program Flow Chart

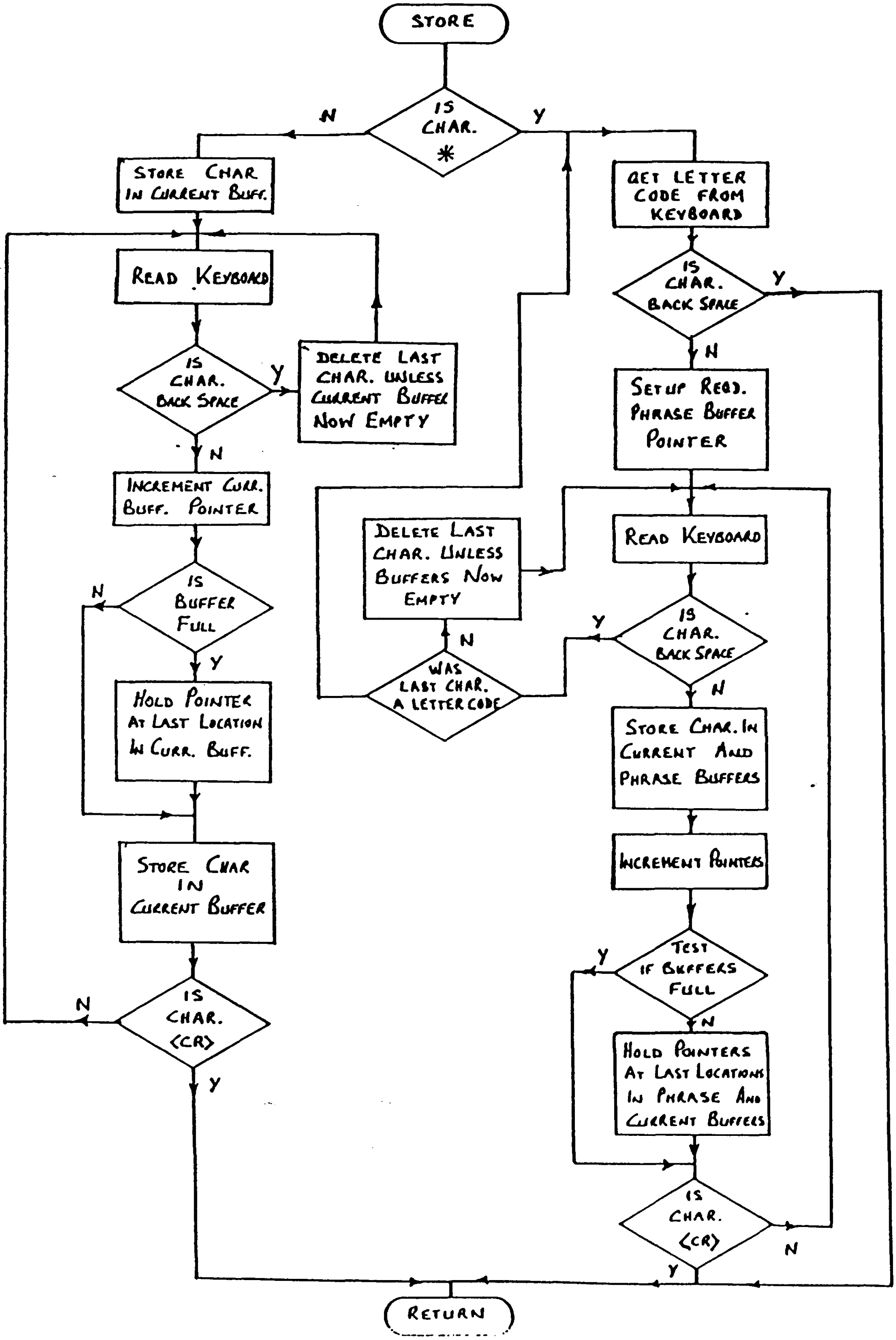

FIGURE A8.6 Text to Speech Mk I: Store Data Flow Chart

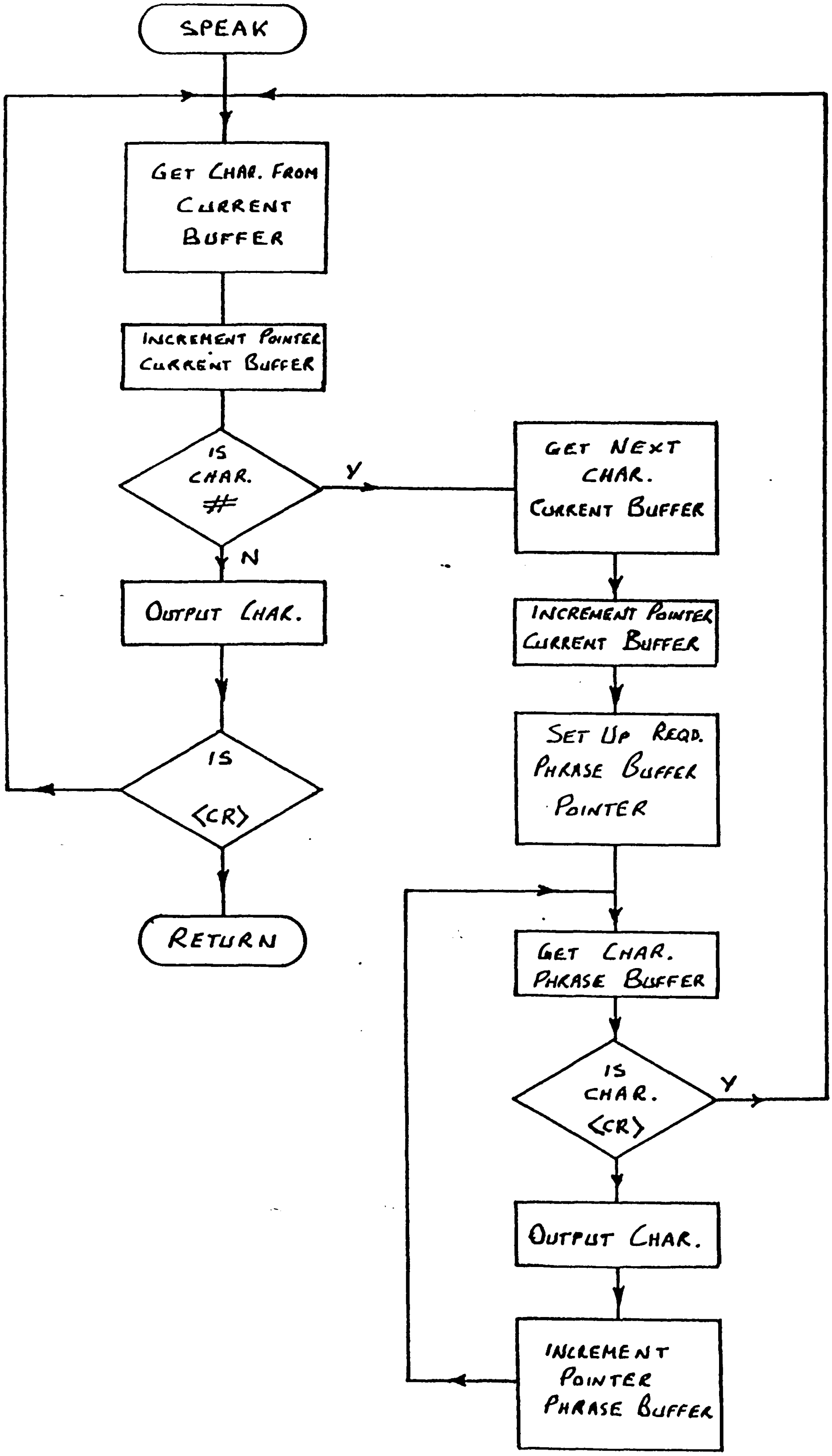

FIGURE A8.7 Text to Speech Mk I: Speak Routine Flow Chart

a) "Back space": The last character entered is deleted from the phrase and/or current buffers.

b) "*": The next entered character is used to obtain a pointer to a phrase buffer.

c) <CR> the entered text string is terminated or if previous character was also <CR> data is transferred from the buffers in the NVR out to the VOTRAX TYPE 'N TALK via the RS232 serial link.

d) Any other character will be stored in the NVR buffers. It is possible to RESET the processor by pressing CTL and RESET simultaneously. This restarts the program and clears the current buffer. All data in the Phrase Buffers in the NVR survives. Power down also leaves this data intact.

The contents on location FFFD (i.e. 17FD) contain the bit pattern of the selected baud rate. The unit is set for a baud rate of 9600. The data format is:- 8 data bits, no parity and two stop bits.

Figures A8.5, A8.6 and A8.7 show the main program, store and speak flow-charts.

Power Supply

The power supply circuit is shown in Figure A8.4. The two PCB mounted transformers give +12 V DC and +12.5 DC unregulated supplies via the bridge rectifiers and smoothing capacitors. The unregulated supplies are fed to voltage regulators to produce the +12 V DC and +5 V DC used on the interfact card. The +12 V unregulated supply and a 9 V AC supply meet the power input requirements of the VOTRAX TYPE 'N TALK.

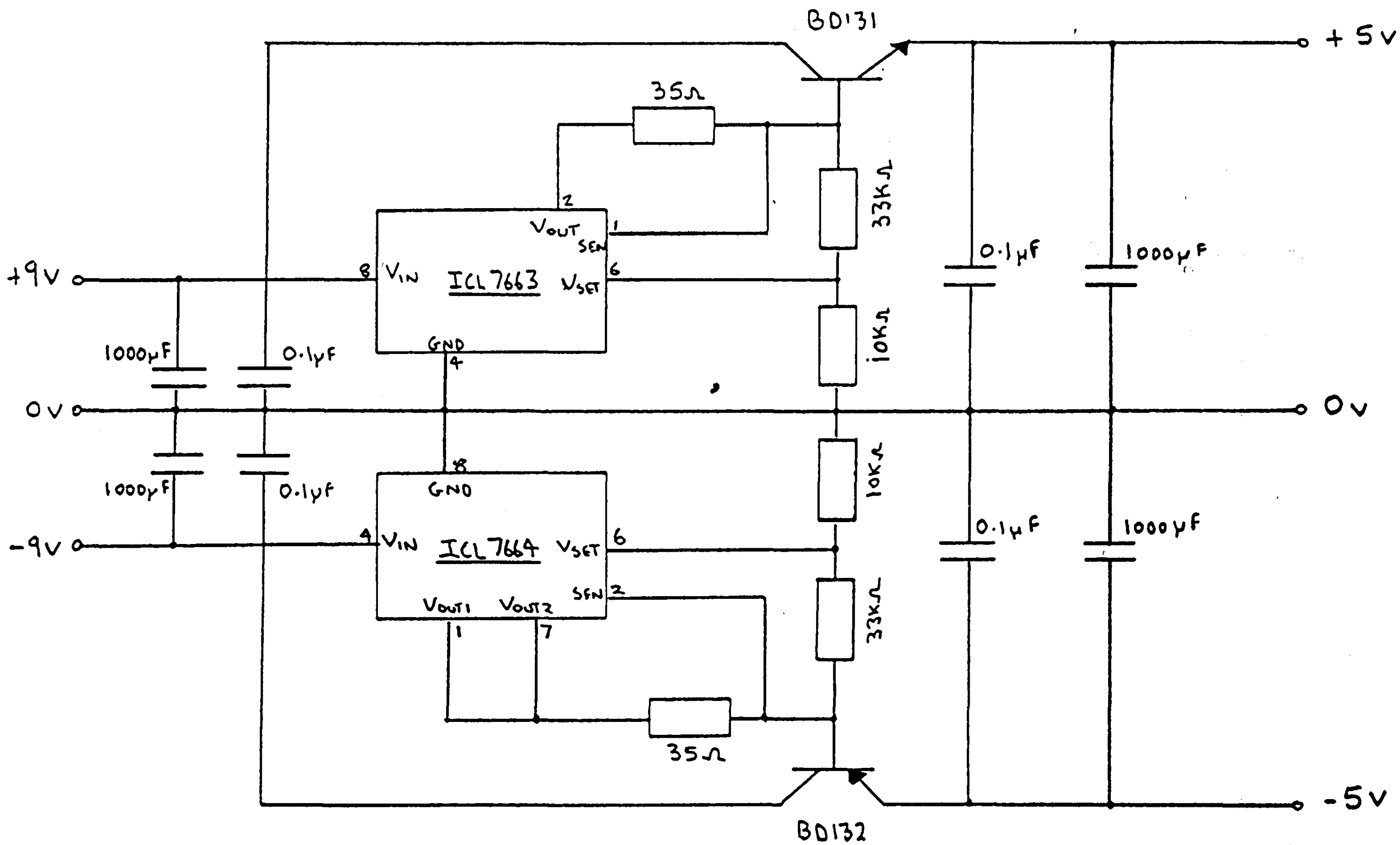

FIGURE A9.1 **The Uvocom: Power Supply**

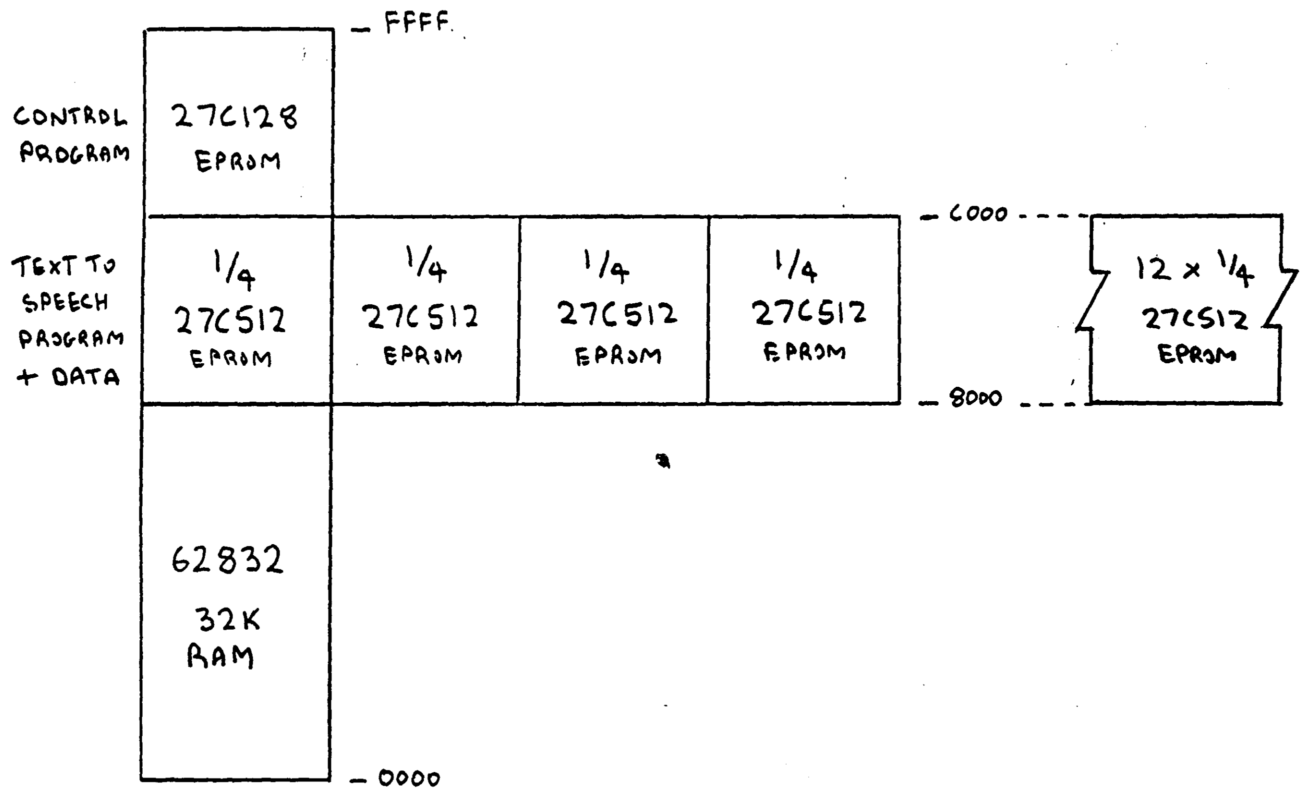

FIGURE A9.2 The Uvocom: Memory Map

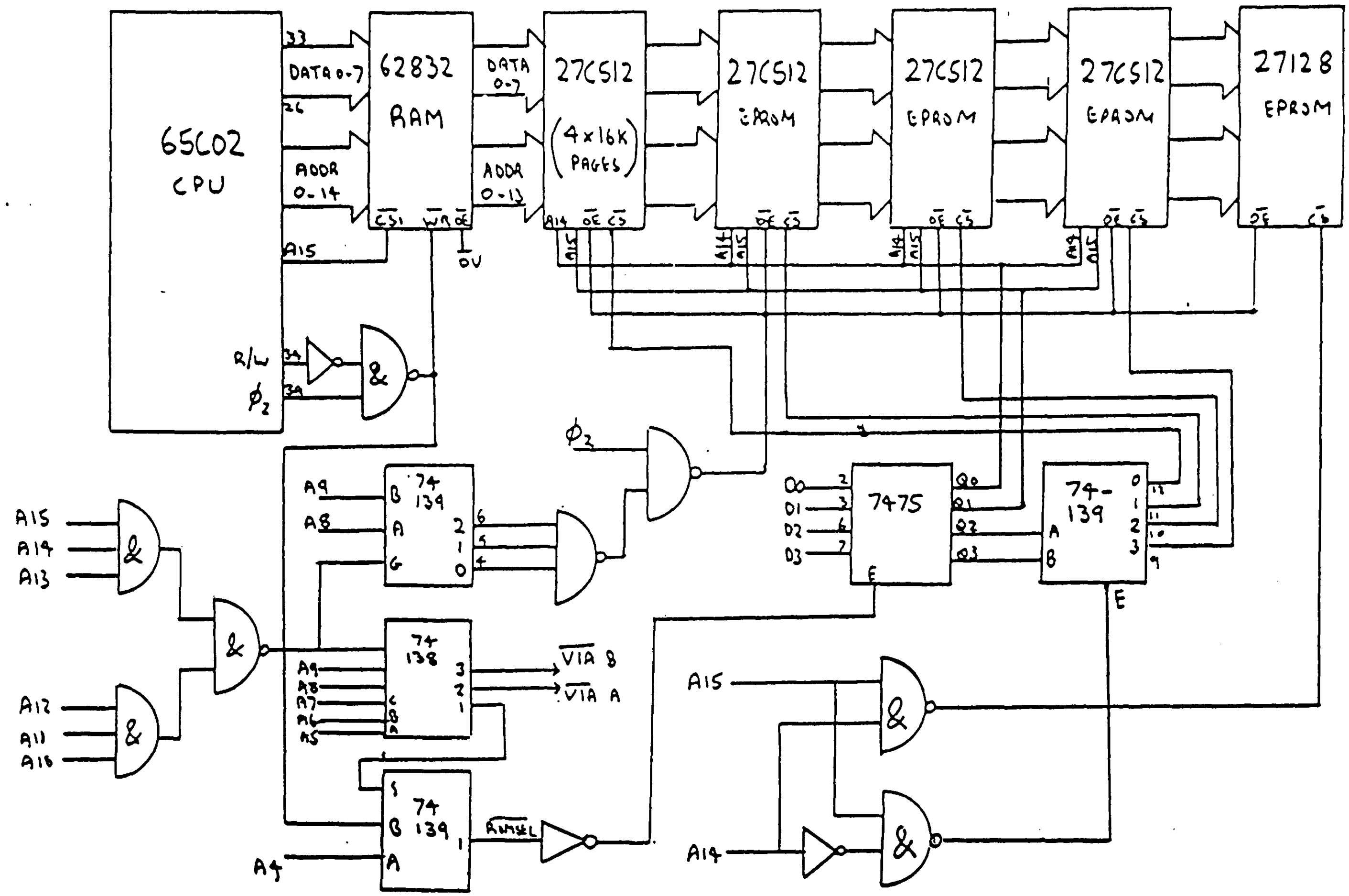

FIGURE A9.3 The Uvocom: Memory Address Decoding

A9.1 Hardware Description

A schematic diagram of the Uvocom is shown in Figure 6.4. It consists of a keyboard, processor, memory, display, speech processor and a speaker.

Power Supply Unit

A circuit diagram of the power supply for the Uvocom is shown in Figure A9.1. Two PP3 9 volt batteries are used as the power source. The 5 volt logic power supply is generated from the 9 volt input voltage via an ICL7663 voltage regulator and transistor driver. The -5 volt power supply for the speech processor chip is generated in a similar way using the ICL7664 voltage regulator.

Memory Addressing

The microprocessor used is the 65C02 which is a standard 8 bit device. This means that the memory has to be paged to address more than 64 Kbytes of data. The memory map of the device is shown in Figure A9.2.

The cicuitry used to perform the address decoding is shown in Figure A9.3. The 32 Kbyte RAM device occupies locations 0000-8000 and is selected by address line A15.

The 16K byte control program EPROM occupies address locations C000-FFFF and is selected by decoding address lines A14 and A15. The EPROM data devices occupy address space 8000-C000. The 64K byte devices must be paged in 16 K byte blocks in order to fit this

KEYBOARD CONNECTOR

PIN	
1	0V
2	$\overline{RST}$
3	1MHz
4	KB $\overline{EN}$
5	PA4
6	PA5
7	PA6
8	PA0
9	PA1
10	PA2
11	PA3
12	PA7
14	CA2
15	+5V

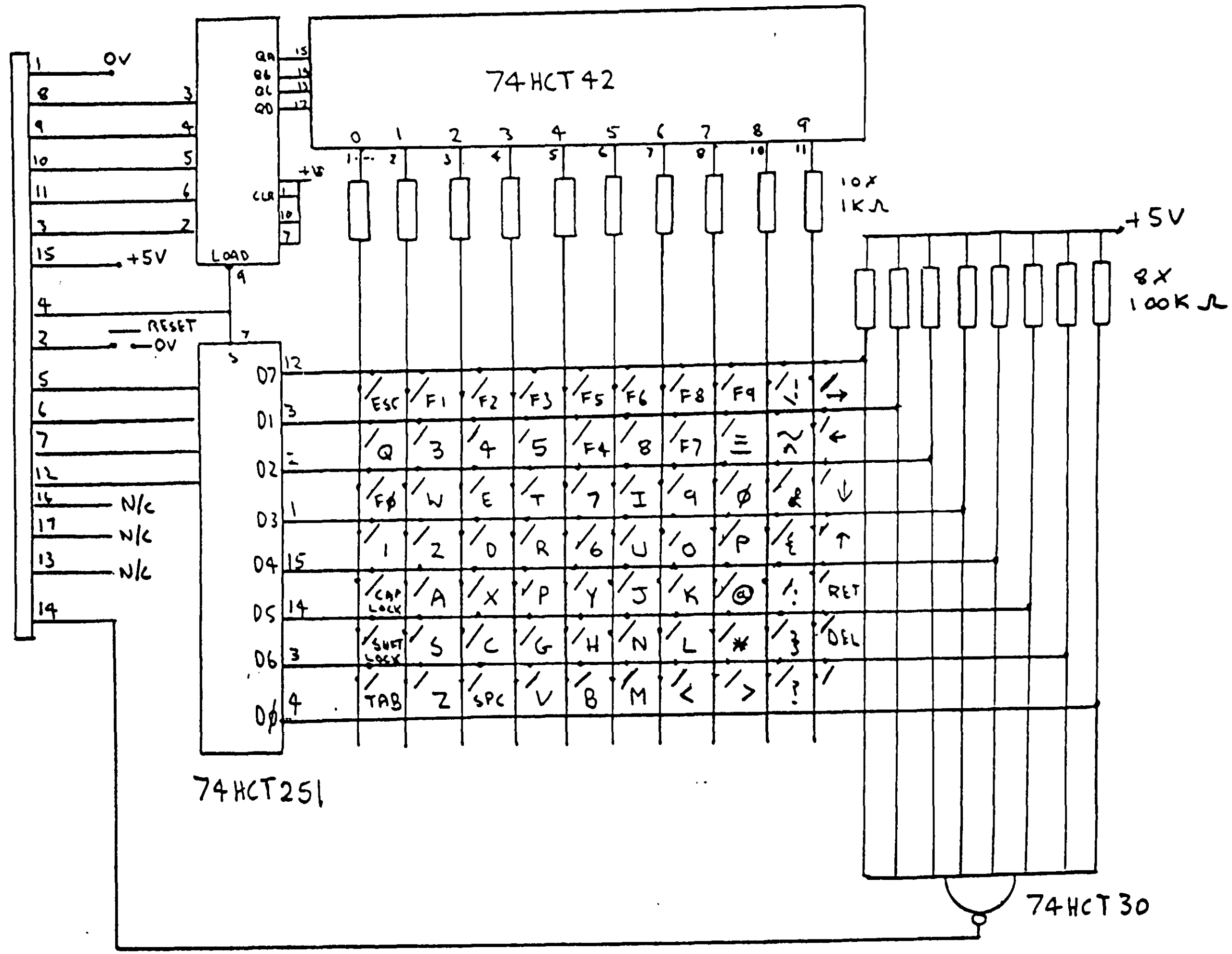

FIGURE A9.4 The Uvocom: Keyboard Circuit

memory space. The 16 Kbyte blocks can be selected from software. This is achieved by writing an EPROM number code (0-15) to an EPROM select register. This register is memory mapped at location &FE30. The register consists of a 74LS75 4 bit latch and a 74LS139 1 of 4 decoder. Data bits D0 and D1 are used by the 74LS75 to select a 16K block within the 64K byte EPROM and bits D2 and D3 are used by the 74LS139 to select 1 of 4 64Kbyte EPROMS.

The Keyboard Circuit

The keyboard circuit is shown in Figure A9.4. and consists of a 10 by 8 matrix of normally open contact switches which are mounted on a metal plate. Connections to the contacts of these switches are made by a printed circuit board.

The keyboard is interfaced to the circuit using the same technique implemented on the BBC microcomputer. A synchronous binary counter 74LS163 is clocked by the 1MHz clock. The outputs from this counter are decoded by a BCD to decimal decoder 74HCT42. The ten outputs from this decoder are connected to the column lines of the keyboard matrix. In this way each column of the keyboard matrix is pulsed low then high. Depression of any key results in the output from an 8 input NAND chip 74HCT30 pulsing high and as the counter pulse passes the column to which the key is connected the output from this chip interrupts the microprocessor using the CA2 line of the 65C22 VIA. On recognition of this interrupt, the microprocessor executes a keyboard reading routine to discover which key was depressed. This is achieved by latching the BCD of each column in turn directly into the 74HCT163 binary counter using outputs PA0 to PA3 of the 65C22 VIA, thus interrogating each column in turn. At the same time outputs

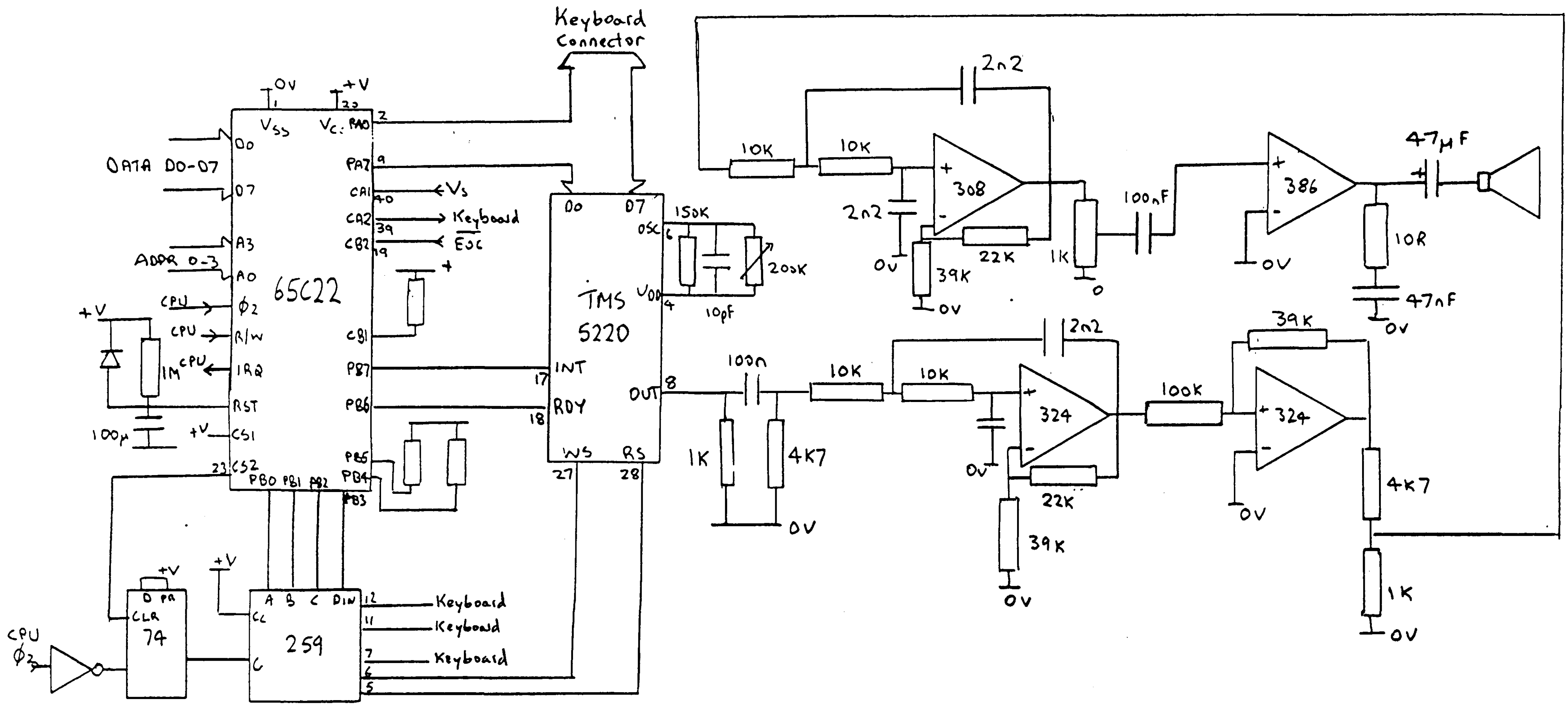

FIGURE A9.5 The Uvocom: Speech Processor and Audio Stage

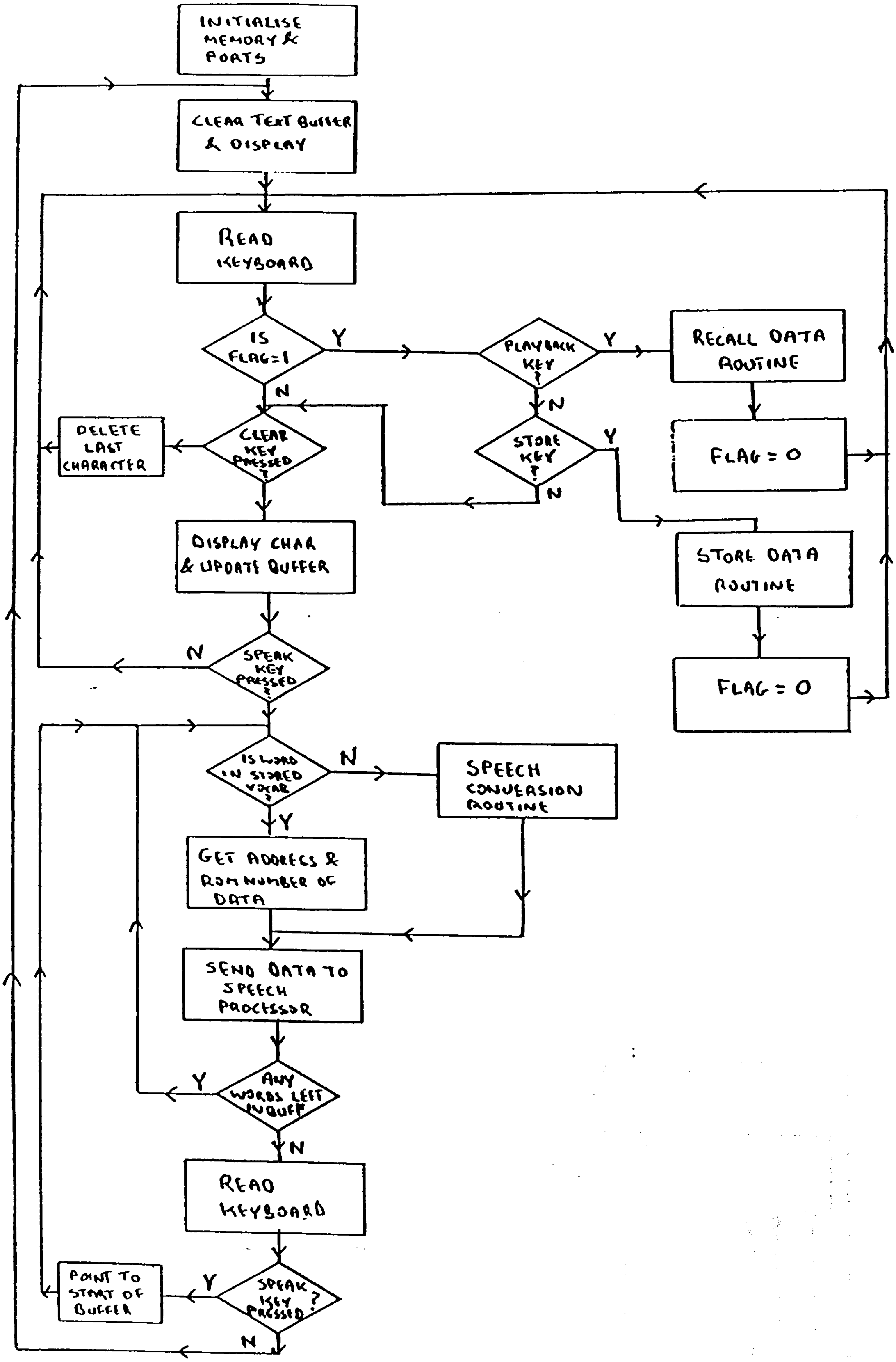

FIGURE A9.6 The Uvocom: Main Program Flow Chart

PA4, 5 and 6 of the VIA are used to load data into the 74HCT251 data selector chip. Each row is selected in turn by the 3 bit codes on PA4,5 and 6. The logic level on a particular row appears at the output of the data selector when selected. In this way the keyboard is scanned for the coincidence created when a key is depressed.

The Speech Processor and Audio Stage

A diagram of the speech processor and audio stage is shown in Figure A9.5.

Speech data for the 5220 chip is taken from the data bus which is provided by port A of the 65C22 VIA.

Interrupt (INT) and ready (RDY) output signals from the speech processor are fed to lines PB6 and PB7 of the VIA respectively. In this way the speech processor is under software control. The analog speech output signal from the 5220 processor is passed through a high pass passive filter to remove any DC signal component. The signal is then passed through a low pass second order active filter with a 5KHz cut-off frequency. The output from the LM308 operational amplifier of the filter is passed via a volume control (internal preset) and AC coupling to the LM386 audio amplifier and to the internally mounted loudspeaker.

A9.2 Software Design

A flow chart of the main program loop is shown in Figure A9.6 On startup the system is initialised by clearing the memory buffers and the display and configuring the input/output ports. A prompt "?" is displayed to inform the user that the device is ready for input of

Vowels		
E - see	i - tin	e - pet
a - cat	u - tub	AR - class
o - dog	OR - fall	oo - put
OO - food	ER - hurt	A - play
I - line	O - bone	OY - boy
AW - clown	EA - steer	ea - pear
UR - tour	UH - utter	

Consonants		
p - pot	B - bad	T - tea
D - dot	K - can	G - get
CH - chest	J - jam	F - fun
V - van	TH - thin	DH - then
S - sad	Z - zone	SH - shed
ZH - vision	H - ham	M - mad
N - nod	NG - sing	L - lip
R - run	Y - yet	W - wet

TABLE A9.1 Phoneme Codes for the Uvocom

A-Rule

```
"/<A> =/AR/"
" <ARE> =/AR R/"
" <AR>O=/AR R/"
"<AR>$=/e R/"
" ^<AS>$=/A S/"
"<A>WA =/AR/"
"<AW>=/o/"
" :<ANY>=/e N E/"
"<A>^+$=/A/"
"$:<ALLY>=/AR L E/"
" <AL>$=/AR L/"
"<AGAIN>=/AR G e N/"
"$<AG>E=/i J/"
"<A>^+$=/a/"
" :<A>^+ =/A/"
"<A>^%=/A/"
"<ARR>=/a R/"
" :<AR> =/AA R/"
"<AR> =/ER/"
"<AR>=/AR R/"
"<AIR>=/A R/"
"<AI>=/A/"
"<AY>=/A/"
"<AU>=/o/"
"$:<AL> =/AR L/"
"$:<ALS> =/AR L Z/"
"<ALK>=/o K/"
"<AL>^=/o L/"
" :<ABLE>=/A B AR L/"
"<ABLE>=/AR B AR L/"
"<ANG>+=/A N J/"
"<A>=/a/"
```

B-Rule

```
"/ <BE>^$=/B i/"
"<BEING>=/B E i NG/"
" <BOTH> =B O TH/"
" <BUS>=/B i Z/"
"<BUIL>=/B i L/"
"<B>=/B/"
```

C-Rule

```
"/^E<CH>=/K/"
"<CH>=/CH/"
" S<CI>$=/S I/"
"<CI>A=/SH/"
"<CI>O=/SH/"
"<CI>EN=/SH/"
"<C>+=/S/"
"<CK>=/K/"
"<COM>%=K u M/"
"<C>=/K/"
```

D-Rule

```
"/$:<DED> =/D i D/"
".E<D> =/D/"
"$^:E<D> =/T/"
" <DE>^$=/D i/"
" <DO> =/D OO/"
" <DOES>=/D u Z/"
" <DOING>=/D OO i NG"
" <DOW>=/D AW/"
"<DU>A=/J OO/"
"<D>=/D/"
```

TABLE A9.2 Text to Phoneme Rules

E-Rule

```
"/$:<E>=//"
" :<E> =/E/"
"$<ED> =/D/"
"$:<E>D =//"
"<EV>ER=/e V/"
"<ERI>$=/E R E/"
"<ERI>=/e R i/"
"$:<ER>=/ER/"
"<ER>$=/e R/"
"<ER>=/ER/"
" <EVEN>=/E V e N/"
"$:<E>W=//"
"£<EW>=/00/"
"<EW>=/Y 00/"
"<E>O=/E/"
"$:&<ES> =/i Z/"
"$:<E>S =//"
"$:<ELY> =/L E/"
"$:<EMENT>=/M e N T/"
"<EFUL>=/F oo L/"
"<EE>=/E/"
"<EARN>=/ER N/"
" <EAR>=/ER/"
"<EAD>=/e D/"
"$:<EA> =/E AR/"
"<EA>SU=/e/"
"<EA>=/E/"
"<EIGH>=/A/"
"<EI>=/E/"
" <EYE>=/I/"
"<EY>=/E/"
"<EU>=/Y 00/"
"<E>=/e/"
```

F-Rule

```
"/<FUL>=/F oo L/"
"<F>=/F/"
```

G-Rule

```
"/<GIV>=/G i V/"
" <G>I^=/G/"
"<GE>T=/G e/"
"SU<GGES>=/G J e S/"
"<GG>=/G/"
" B$<G>=/G/"
"<G>+=/J/"
"<GREAT>=/G R A T/"
"$<GH>=//"
"<G>=/G/"
```

H-Rule

```
"/ <HAV>=/H a V/"
" <HERE>=/H E R/"
" <HOUR>=/AW R/"
"<HOW>=/H AW/"
"<H>$=/H/"
"<H>=//"
```

TABLE A9.2 Text to Phoneme Rules

I-Rule

```
"/ <IN>=/i N/"
" <I> =/I/"
"<IND>=/I N D/"
"<IER>=/E ER/"
"$:R<IED> =/E D/"
"<IED> =/I D/"
"<IEN>=/E e N/"
"<IE>T=/I e/"
" :<I>%=/I/"
"<I>%=/E/"
"<IE>=/E/"
"<IR>$=/I R/"
"<IZ>%=/ I Z/"
"<IS>%=/I Z/"
"<I>D%=/I/"
"+^<I>^+=/i/"
"<I>T%=/I/"
"$^:<I>^+=/i/"
"<I>^+=/I/"
"<IR>=/ER/"
"<IGH>=/I/"
"<ILD>=/I L D/"
"<IGN> =/I N/"
"<IGN>^=I N/"
"<IGN>%=/I N/"
"<IQUE>=/E K/"
"<I>=/i/"
```

J-Rule

```
"/<J>=/J/"
```

K-Rule

```
" <K>N=//"
"<K>=/K/"
```

L-Rule

```
"/<LO>C$=/L O/"
"L<L>=//"
"$^:<L>%=/AR L/"
"<LEAD>=/L E D/"
"<L>L=/L/"
```

M-Rule

```
"/<MOV>=/M OO V/"
"<M>=/M/"
```

N-Rule

```
"/E<NG>+=/N J/"
"<NG>R=/NG G/"
"<NG>$=/NG G/"
"<NGL>%=/NG G AR L/"
"<NG>=/NG/"
"<NK>=/NG K/"
" <NOW> =/N AW/"
"<N>=/N/"
```

TABLE A9.2 Text to Phoneme Rules

O-Rule

```
"/<OF> =/o V/"
"<OROUGH>=/ER O/"
"$<OR> =/ER/"
"$:<ORS> =/ER Z/"
"<OR>=/o R/"
" <ONE>=/W u N/"
"<OW>=/O/"
" <OVER>=/O V ER/"
"<OV>=/u V/"
"<O>^%=/O/"
"<O>^EN=/O/"
"<O>^I$=/O/"
"<OL>D=/O L/"
"<OUGHT>=/o T/"
"<OUGH>=/u F/"
" <OU>=/AW/"
"H<OU>S$=/AW/"
"<OUS>=/AR S/"
"<OUR>=/o R/"
"<OULD>=/oo D/"
" <OU>^L=/u/"
"<OUP>=/OO P/"
"<OU>=/AW/"
"<OU>=/OY/"
"<OING>=/O i NG/"
"<OI>=/OY/"
"<OOR>=/o R/"
"<OOK>=/oo K/"
"<OOD>=/oo D/"
"<OO>=/OO/"
"<O>E=/O/"
"<O> =/O/"
"<OA>=/O/"
" <ONLY>=O N L E/"
" <ONCE>=/W u N S/"
"C<O>N=/AR/"
"<O>NG=/o/"
" ^:<O>N=/u/"
"I<ON>=/AR N/"
"$:<ON> =/AR N/"
"$^<ON>=/AR N/"
"<O>ST =/O/"
"<OF>^=/o F/"
"<OTHER>=/u DH ER/"
"<OSS> =/o S/"
"$^:<OM>=/u M/"
"<O>=/AR/"
```

P-Rule

```
"/<PH>=/F/"
"<PEOP>=/P E P/"
"<POW>=P AW/"
"<PUT>=/P oo T/"
"<P>=/P/"
```

Q-Rule

```
"/<QUAR>=/K W o R/"
"<QU>=K W/"
"<Q>=/K/"
```

R-Rule

```
"/ <RE>$=/R i/"
"<R>=/R/"
```

S-Rule

```
"/<SH>=/SH/"
"$<SION>=/ZH o N/"
"<SOME>=/S u M/"
"$<SUR>$=/AH ER/"
"<SUR>$=/SH ER/"
"$<SU>$=/AH OO/"
"$<SSU>$=/SH OO/"
"$<SED> =Z D/"
"$<S>$=/Z/"
"<SAID>=/S e D/"
"^<SION>=/SH o N/"
"<S>S=//"
".<S> =/Z/"
"$:.E<S> =/Z/"
"$^:$$<S> =/Z/"
"$^:$<S> =/S/"
"U<S> =/S/"
" :$<S> =/Z/"
" <SCH>=/S K/"
"<S>C+=//"
"$<SM>=/A M/"
"<S>=/S/"
```

TABLE A9.2 Text to Phoneme Rules

T-Rule

```
"/ <THE> =/DH u/"
"<TO> =/T 00/"
"<THAT> =/DH a T/"
" <THIS> =DH i S/"
" <THEY>=/DH A/"
" <THERE>=/DH e R/"
"<THER>=/DH ER/"
"<THEIR>=/DH e R/"
" <THAN> =/DH a N/"
" <THEM> =/DH e M/"
"<THESE> =/DH E Z/"
" <THEM>=/DH e N/"
"<THROUGH>=/TH R 00/"
"<THOSE>=/DH O Z/"
"<THOUGH>=/DH O/"
" <THUS>=/DH u S/"
"<TH>=/TH/"
"$:<TED> =/T i D/"
"S<TI>$N=/CH/"
"<TI>O=/SH/"
"<TI>A=/SH/"
"<TIEN>=/SH o N/"
"<TUR>$=/CH ER/"
"<TU>A=/CH 00/"
" <TWO>=/T 00/"
"<T>=/T/"
```

U-Rule

```
" <UN>I=/Y 00 N/"
" <UN>=/u N/"
" <UPON>=/u P o N/"
"£<UR>$=/oo R/"
"<UR>$=/Y oo R/"
"<UR>=/ER/"
"<U>^ =/u/"
"<U>^^=/u/"
"<UY>=/I/"
" G<U>$=//"
"G<U>%=//"
"G<U>$=/W/"
"$N<U>=/Y 00/"
"£<U>=/00/"
"<U>=Y 00/"
```

V-Rule

```
"/<VIEW>=/V Y 00/"
"<V>=/V/"
```

W-Rule

```
"/ <WERE>=/W ER/"
"<WA>S=/W AR/"
"<WA>T=/W AR/"
"<WHERE>=/W H e R/"
"<WHAT>=/W H o T/"
"<WHOL>=/H O L/"
"<WHO>=/H 00/"
"<WH>=/W H/"
"<WAR>=/W o R/"
"<WOR>^=/W ER/"
"<WR>=/R/"
"<W>=/W/"
```

TABLE A9.2 Text to Phoneme Rules

X-Rule

```
"/<X>=/K S/"
```

Y-Rule

```
"/YOUNG>=/Y u NG/"
" <YOU>=/Y OO/"
" <YES>=/ Y e S/"
" <Y>=/Y/"
"$^:<Y> =/E/"
"$^:<Y>I=/E/"
" :<Y> =/I/"
" :<Y>$=/I/"
" :<Y>^+:$=/i/"
" :<Y>^$=/I/"
"<Y>=/i/"
```

Z-Rule

```
"/<Z>=/Z/"
```

TABLE A9.2 Text to Phoneme Rules

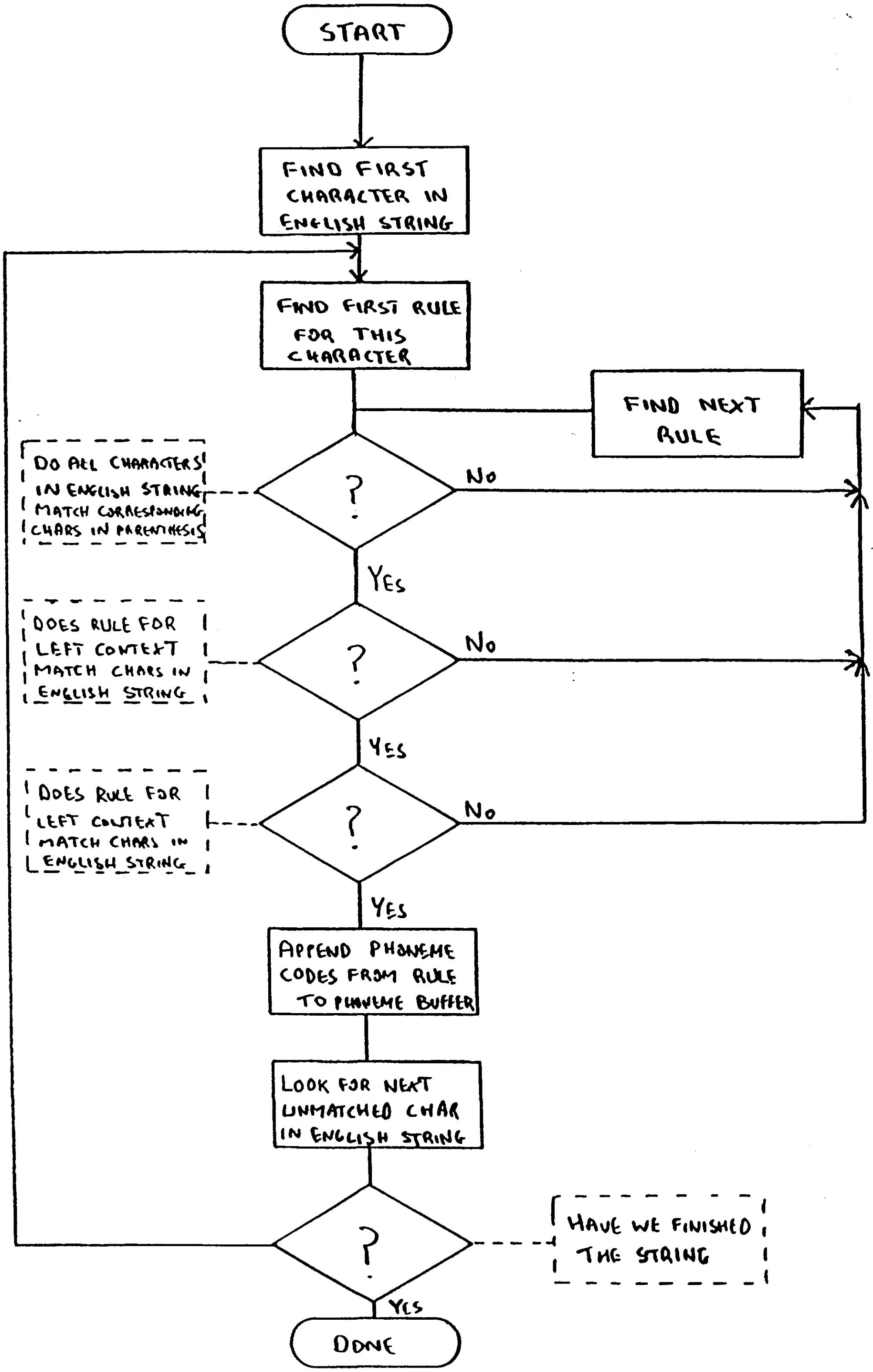

FIGURE A9.7 Text-to-Phoneme Program Flow Chart text. The software monitors the keyboard for a key press and when a key press is detected the program displays the character.

On receiving a 'SPEAK' command from the keyboard the first word in the text buffer is checked to see if it is stored in the core vocabulary. If it is then the address of the speech data and the EPROM number is obtained from the core vocabulary jump table. This information is used to retrieve the speech data from the paged memory and send it to the speech processor. If the word is not stored in the core vocabulary then the text to phoneme conversion routine is called.

Software was written to convert text into the phoneme codes required by the commercial firmware. These phoneme codes are shown in Table A9.1. Letter-to-sound rules for automatic translation of English text into phonetics have been developed by Elovitz (59). These rules were developed for the Votrax device but are easily adapted for any other phomeme input device. Elovitz has shown these rules to produce correct pronounciation for approximately 90% of the words in a list of 8,000 most frequently occuring words.

Each of the 329 letter-to-sound rules of Table A9.2 supplies a pronounciation for the text string enclosed in the angled brackets. Each text string contained in the angled brackets may also have a right context and a left context which must also be examined.

The algorithm used for interpreting the rules is based on an article by Ciarcia (60) and is shown in Figure A9.7.

The program recognises the first character of the input text string and skips down the list to the first applicable rule and attempts to match the rule's angled bracket string to the input text. If there is no match, the process is repeated with the next rule applicable for the letter. If there is a match on the angled bracket

SYMBOL	MEANING
$	One or more vowels*
.	One of B, D, B, G, J, L, M, N, R, W, Z: a voiced consonant
%	One of ER, E, ES, ED, ING, ELY: a suffix
&	One of S, C, G. Z, X, J, CH, SH: a sibilant
^	One consonant**
£	One of T, S, R, D, L, Z, N, J, TH, CH, SH: a consonant influencing the sound of the following u (i.e. rule, mule)
+	One of E, I, Y: a front vowel

* Vowels are A, E, I, O, U, Y
** Consonants are B, C, D, F, G, H, J, K, L, M, N, P, Q, R, S, T, U, V, W, X, Z

TABLE A9.3 Uvocom Control Codes

string, an attempt is made to match first the left and then the right context. If either context match fails, the program proceeds to the next rule. The final rule for each letter contains no context characters thus guaranteeing an eventual match for any letter of the input text.

During an attempt to match a character string, the program may encounter a special symbol in the rule expression. In such a case, the symbol is looked up in a table in memory, and the corresponding subroutine is called. The special control symbols are described in Table A9.3.

Once a match has been achieved, the phoneme codes invoked by the rule are transferred to a phoneme buffer.

The phoneme codes are sent from this buffer to the commercial Phoneme-LPC EPROM which sends the LPC code to the speech processor. The program then returns to the keyboard read routine of the main program loop.

The flow charts of Figure A8.6 and A8.7 apply when the device is used in the recall mode.

SPEECH SYNTHESIS PROJECT - PUBLICATIONS

1. Evans, A. L., Gowdie, R. A., Keating, D., Smith, D. C., Wyper, D. J. and Cunningham, E. A Versatile Speech Output Communication Aid. Journal of Medical Engineering & Technology, 9, 180 - 182 (1985).

2. Keating, D., Evans, A. L., Wyper, D. J. and Cunningham, E. Experience with Low Cost Portable Speech Aids Using LPC Synthesis. HPA Conference on Computers and the Handicapped - Rehabilitation, Assessment, Aids for Living, London. HPA Bulletin, March, p. 10 (1986).

3. Keating, D., Evans, A. L., Wyper, D. J., Bone, I. and Cunningham, E. A Comparison of the Intelligibility of some Low Cost Speech Synthesis Devices. British Journal of Disorders of Communication, 21, pp 167 - 172 (1986).

4. Keating, D., Evans, A. L., Smith, D. C., Wyper, D. J., Cunningham, E. and Bone, I. Text to Speech Using an LPC Synthesiser. Proc. 4th ISAAC Conference, Cardiff. Augmentative and Alternative Communication, 2 (3) p. 80 (1986).

5. Wyper, D. J., Cunningham, E., Keating, D. and Evans, A. L. The Art of Conversation Prompting. Proc. 4th ISAAC Conference, Cardiff. Augmentative and Alternative Communication, 2 (3), p. 124 (1986).

6. Wyper, D. J., Evans. A. L., Cunningham, E. and Keating, D. A Communication Aid for the Non Vocal in the Intensive Care Unit. Proc. 4th ISAAC Conference, Cardiff. Augmentative and Alternative Communication, 2 (3), p. 123 - 4 (1986).

7. Keating, D., Evans, A. L., Gowdie, R. A., Wyper, D. J., Cunningham, E. and Bone I. The Pocket Speech Aid - A Versatile Tool. Proc. 4th ISAAC Conference, Cardiff, Augmentative and Alternative Communication, 2, (4), p. 148 (1986).

8. Keating, D., Evans, A. L., Wyper, D. J., Cunningham, E. and Bone, I. Synthetic Speech Aids for the Disabled. proc. 43rd Annual HPA Conference, Coventry. 1986. Clinical Physics and Physiological Measurement, 8 (1), p. 87 (1987).

9. Keating, D. Communicating with the Pocket Speech Aid. Communicating Together. 5 (1) p. 10 (1987).

10. Keating, D. and Wyper, D. J. Low Cost Portable Speech Aids for the Disabled. Linguistics Lecture, Edinburgh University (1987).

REFERENCES

1. Enderby, P. (1986). Relationships between dysarthric groups. British Journal of Disorders of Communication, 21, 189-197.

2. Bloodstein (1979). Speech Pathology: An Introduction. Houghton Mafflin Company, USA.

3. Darley, F.L., Aronston, E. and Brown, M.D. (1985). Motor Speech Disorders. Saunders, Philadelphia.

4. Rosenbeck, J.C. (1976). Treatment of Apraxia of Speech. Proceedings of the Acoustics Speech and Hearing Association, Houston.

5. Enderby, P. and Philip, R. (1986). Speech and Language Handicap: towards knowing the size of the problem. British Journal of Disorders of Communication, 21, 151-165.

6. Gibberd, F.B., Oxtoby, M. and Jewell, F.P. (1985). The treatment of Parkinson`s Disease: a consumers view. Health Trends, 17, 19-21.

7. Meritt, H. (1979). A textbook of Clinical Neurology. Lea and Fabiger, Philadelphia.

8. Darley, F., Brown, J. and Goldstein, N. (1972). Dysarthria in Multiple Sclerosis. Journal of Speech and Hearing Research, 15, 229-245.

9. Mcalpine, D., Compton, N. and Lumsden, C. (1955). Multiple Sclerosis, Livingstone, London.

10. Harding, A.E. and Zikha, K.J. (1981). Pseudo dominant inheritance in Fredreich's Ataxia. Journal of Medical Genetics. 18, 285-287.

11. Walton, J. (1977). Brain`s Diseases of the Nervous System. Churchill Livingstone, London.

12. Office of Health Economics (1980). Huntington`s Chorea, 67. White Cross Press, London.

13. Kurland, L.T. (1977). Epidemiology of Amyotrophic Lateral Sclerosis with emphasis on antecedent events from case control comparisons. In: Motor Neuron Disease, Ross, F.C. (Ed). Grune & Stratton, New York.

14. Garfinkle, T.J. & Kimmelman, C.P. (1982). Neurologic Disorders: Anyotrophic Lateral Sclerosis, Mayashenia Gravis, Multiple Sclerosis and Poliomyelitis. American Journal of Otolaryngology, 14, 204-212.

15. Carter, A.B. (1964). Cerebral Infarction. MacMillan, New York.

16. Marquardsen, J. (1969). The natural history of cerebrovascular disease. Acta Neurologica Scandanavia Supplement, 45, 90-188.

17. Field Research Division of DHSS. (1975). Epedemiology of Head Injuries in England and Wales. DHSS, London.

18. Nolan, C. (1981). Dam-burst of dreams. Weiderfield & Nicholson, London.

19. Bristow, G. (1984). Electronic Speech Synthesis. Granada, London.

20. Witten, I.H. (1986). Making Computers Talk. Prentice Hall, New Jersey.

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

22. Shannon, C.E. and Weaver, W. (1963). The Mathemmatical Theory of Communication. University of Illinois Press. Illinois.

23. Black, H.S. (1953). Modulation Theory. Van Nostrand, New York.

24. Holmes, J.N. (1982). A survey of methods for digitally encoding speech signals. The Radio and Electronic Engineer, 52 (6), 267-276.

25. Schindler, H.R. (1970). Delta Modulation. IEEE Spectrum, October, p69.

26. Panter, P.F. (1965). Modulation, Noise and Spectral Analysis. McGraw Hill. New York.

27. OKI Semiconductor Application Note. (1982). Sumiltaneous speech analysis and synthesis with MSM5218.

28. Evans, A.L. and Fenner, J. (1987). Random-access phrase recorder. Electronics and Wireless World, 92, (1611), 54-56.

29. Markel, J.D. and Gray, A.H. Jnr. (1976). Linear Prediction of Speech. Springer Varlag, New York.

30. Saito, S., Itakura, F. (1966). The Theoretical Consideration of Statistically Optimum Methods for Speech Spectral Density. Report No. 3107, Electrical Communication Laboratory, N.T.T., Tokyo.

31. Atal, B.S. and Schroeder, M.R. (1987). Predictive Coding of Speech Signals. Proc 1967 Conf. on Communication and Process., 360-361.

32. Wiener, N. (1966). Extrapolation, Interpolation and Smoothing of Stationary Time Series. M.I.T. Press, Cambridge, Massachusetts.

33. Bristow, G.J. and Gunwardana, R. (1983). A new method of speech data preparation using a portable microcomputer system. Proceedings of IEEE Colloquium, London.

34. Evans, A.L., Gowdie, R.A., Keating, D., Smith, D.C., Wyper, D.J. and Cunningham, E. (1985). A versatile speech output communication aid. Journal of Medical Engineering and Technology, 9, 180-182.

35. Voiers, W.D. (1977). Diagnostic evaluation of speech intelligibility. In: Speech Intelligibility and Speaker Recognition, 374-387. Hawley, M.E. (Ed). Dowden, Hutchison and Ross, Stroudsberg, P.A.

36. Miller, G.A. and Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Societ of America, 27, 338-352.

37. Howes, D. (1966). A word count of spoken English. Journal of Verbal Learning and Verbal Behaviour, 5, 572-604.

38. Hellen, P.A.T. (1985). Military secure speech using HF. Electronics and Power, March 232-238.

39. Kraat, A.W. (1985). Communication interaction between aided and natural speakers, Canadian Rehabilitation Council for the Disabled, Toronto.

40. Foulds, R.A. (1980). Communication rates for non-speech expression as a function of manual tasks and linguistic constraints. Proceedings of The 3rd International Conference on Rehabilitation Engineering, Toronto.

41. Beukelman, D.R., Yorkston, K.M., Poblette, M., Naranjo, C. (1984). Frequency of word occurrence in communication samples produced by adult communication aid users. Journal of Speech and Hearing Disorders, 49, 360-367.

42. Shane, H.C. (1980). Approaches to Assessing the Communication of Non-Oral Persons. In: Nonspeech Language and Communication: Analysis and Intervention, 198-224. Schiefelbush, E. (Ed). University Park Press, Baltimore.

43. Goodenough-Trepagnier, C. and Galdieri, B. (1984). Slow message production rate and receivers` impatience. Proceedings of The 2nd International Conference on Rehabilitation Engineering, Ottowa.

44. Smith, W. and Crook, S.B. (1981). Phonemes, allophones and LPC team to synthesize speech. Electronic Design. 25 June, 121-127.

45. Schonell, F.J. and Meddleton, I.G. (1956). A study of the oral vocabulary of adults. University of London Press, London.

46. Minneman, S.L. (1986). Keyboard optimisation technique to improve output rate of disabled individuals. Proceedings of the 9th Annual RESNA Conference, Minnesota.

47. Damper, R.I., Burnett, J.W., Gray, P.W., Straus, L.P. and Symes, R.A. (1987). Hand held text-to-speech device for the non-vocal disabled. Journal of Biomedical Engineering, 9, 332-340.

48. Arnott, J.L., Pickering, J.A., Swiffin, A.L. and Battison, M.A. (1984). An adaptive and predictive communication aid for the disabled exploits the redundancy in natural language. Proceedings of the 2nd International Conference on Rehabilitation Engineering, Ottawa.

49. Clark, P., Evans, A.L. and Tremewan, R. (1986). Single switch use of a portable computer. Medical and Biological Engineering and Computing. 25, 115-117.

50. Goodenough-Trepagnier, C., Rosen, M.J. and Galdieri, B. (1986). Word menu reduces communication rate. Proceedings of the 9th Annual RESNA Conference, Minnesota.

51. Soede, M. and Foulds, R.A. (1986). Dilemma of prediction in communication aids and mental load. Proceedings of the 9th Annual RESNA Conference, Minnesota.

52. Rosen, M.J. and Goodenough-Trepagnier, C. (1983). Development of a computer-aided system for prescription of non vocal devices: a progress report. 6th Annual Conference on Rehabilitation Engineering, San Diego.

53. Skinner, C., Wirz, S. and Thomson, I. (1984). The Edinburgh functional communication profile. Winslow Press

54. Le Cardinal, C. (1986). Systemic model of interpersonal communication applied to the communication problems of the handicapped. Proceedings of the 4th ISAAC Conference, Cardiff.

55. Foulds (1986). Human quality synthetic speech using LPC diphones. Proceedings of the 4th ISAAC Conference, Cardiff.

56. Orkin, M. and Drogin, R. (1977). Vital Statistics. McGraw Hill, New Delhi.

57. Wonnacott, T.H. and Wonnacott, R.J. (1972). Introductory Statistics. Wiley, Toronto.

58. Bradley, R.A. (1984). Paired Comparisons: Some Basic Procedures and Examples. In Handbook of statistics Vol. 4, Kirishnaiah, P.R. and Sen, P.K. (Eds). Elsevier Science.

59. Elovitz, H.Z., Johnson, R., McHugh, A. and Shore, J.E. (1976). Letter-to-sound Rules for the Automatic Translation of English Text to Phonetics. IEEE Trans on ASSP, 24, 6, 446-459.

60. Ciarcia, S. (1982). Build the Microvox Text-to-Speech Synthesiser Part 2: Software. Byte, October 40-64.