PERSONALISED ADPCM SPEECH PLAYBACK DEVICES FOR THE NON-VOCAL AND THE NON-VOCAL PHYSICALLY HANDICAPPED

R. B. Reilly & A. M. DePaor

Abstract:

Two personalised instantaneous speech playback devices were designed and developed. Both centred upon the memory efficiency advantages of Adaptive Differential Pulse Coded Modulation (ADPCM). One system is handheld and offers 64 seconds of speech. The second system hasdictionary of 230 four second phrases and is PC based.

Medical Assessment :

The problem of communication for the vocally handicapped is of grave concern. The communication aids currently available for this group of society include text-to-speech synthesizers and portable LPC coders (ref. 1 & 2). The use of these devices is restrictive due to poor quality speech and the use normally of a mid atlantic accent. Some offer good quality speech although are somewhat restrictive in use (ref. 3). The non-vocal physically handicapped also have few systems available to them for communication. The severity of their disabilities normally limits their use such systems. Following a study conducted by the National Medical Rehabilitation Centre it became evident that a syntheziser system would have to preserve quality accent and sex. Also apparent from this study was the fact that full text-to-speech was not required by the majority of potential users but instead short well spoken phrases. Cost also had a bearing on the communication aid prescribed. An inexpensive instrument would provide a more equitable therapy service.

Handheld Communicator:

A small ADPCM portable playback device was developed. The electronics of the device was centred on the CMOS OKI Semiconductor MSM5218 speech processor (ref. 4). Using this chip speech is automatically modulated from 12 to 4 bits and stored. The recording circuitry is linked to IBM PC via an I/O interface card. The quantity of speech stored is limited only by the fixed disk capacity of the PC. The speech is input from a microphone, filtered, digitized and compressed.

The same speech processing chip is incorporated in the playback unit. The handheld unit also consists of small single chip microcontroller CMOS NEC uPD78C10. A 4 x 4 keyboard is provided to allow access to 16 four second phrases (Memory of 2M bytes). The controller scans the keyboard and selects the memory locations required. Each nibble present at each memory address contains one sample of speech data. The data is expanded and output via a filter and audio amplifier.

The use of a keyboard by sufferers of Cerebal Palsy can be awkward. For such users an external switch selection mechanism has been provided. The controller scans each of the four columns in sequence illuminating small LEDs in the process. A positive action on the switch causes the column

Dept of Electronic and Electrical Engineering, University College Dublin, Belfield, Dublin4, Ireland

corresponding to be selected. The controller then scans each row. A positive action on the switch results in an individual key being selected. Such a switch can be a chin-bar, a suck-or-blow switch or an ultrasonic proximity switch mounted on a wheelchair or bedframe. Keypad icons provide the user with a reminder of the phrase uttered at that key location. Different icons can be inserted behind each key as different phrases are recorded. The length of individual phrases can be altered allowing variable-length messages according to the needs of the user. The allocation of the 16 different key areas can be changed to form 4 or only 2 areas. This allows those restricted in dexterity to access an area as opposed to one key location.

PC Based Communicator:

As the handheld communicator was designed for those who which to use a limited number of phrases in everyday situations, a PC based aid was developed for those who which to use a more sophisticated communicator. The system was designed for use with non-vocal quadraplegics in mind. The communicator uses two switches which are continually monitored. The screen initially displays the first letter of the alphabet four times normal size in the top left hand corner. The use of one switch allows the alphabet to be scrolled. The second switch selects a letter. That letter once chosen results in a dictionary of words starting with that letter being displayed on the screen. The dictionary contains the words most widely used by the patient. The patient can scroll through the dictionary using the first switch and select a word using the second.

Accompanying the words in the dictionary are ADPCM encoded versions of these words. Upon building a sentence the selection of *Speak* from any dictionary window will automatically output the required stored speech data in RAM to circuitry containing the MSM5218 speech processor and an external audio amplifier and loudspeaker. The full sentence will be spoken. The number of words presently stored in the dictionary is approximately 230. This number can be increased although it was found that no more than 200 are regularly used.

Use of Communication Aids:

The supervision of these communication aids is the responsibility of the speech therapists in the medical location at which they are being used. The choice of phrases for the handheld playback unit is directed by the therapist. The voice chosen is normally that of an individual similar in sex,dialect and age. The recording of voices of volunteers is carried out by the therapist. The encoded phrases can be replayed and recorded easily before they are downloaded to the playback unit or the dictionary based communicator. The storage within the handheld unit currently consists of EPROM however a RAM version is being developed. For use with the PC based communicator with quadraplegics, the type and the location of the switches is vital. A small training routine is provided to aid in the alignment of the switches. It simultaneously provides instruction to the user on the action of each switch. The use of speech with this communicator has made people within the vacinity of the patient more attentive, compared with that of a screen textwriter.

Conclusion:

The adaptablity provided by these communication systems has dramatically improved the interaction between the user and his community. The two speech message devices offer versatile systems personalised to a neglected group within society. The possibility of affordable instantaneous speech communication can increase the value of their lives beyond measure.

References:

1. Speech Technology (Ireland) Limited 1989 "Acorn : The Portable Speech Communicator".
2. Lancaster University's Aura Speech System
3. D.Keating, D.C.Smith, A.L.Evans, D.J.Wyper, E.Cunningham "The Pocket Speech Aid", Medical & Biological Engineering & Computing, May 1989, Vol 27.
4. OKI Semiconductors Application Note 1988 "Simultaneous Speech Analysis and Synthesis with the OKI MSM5218 "

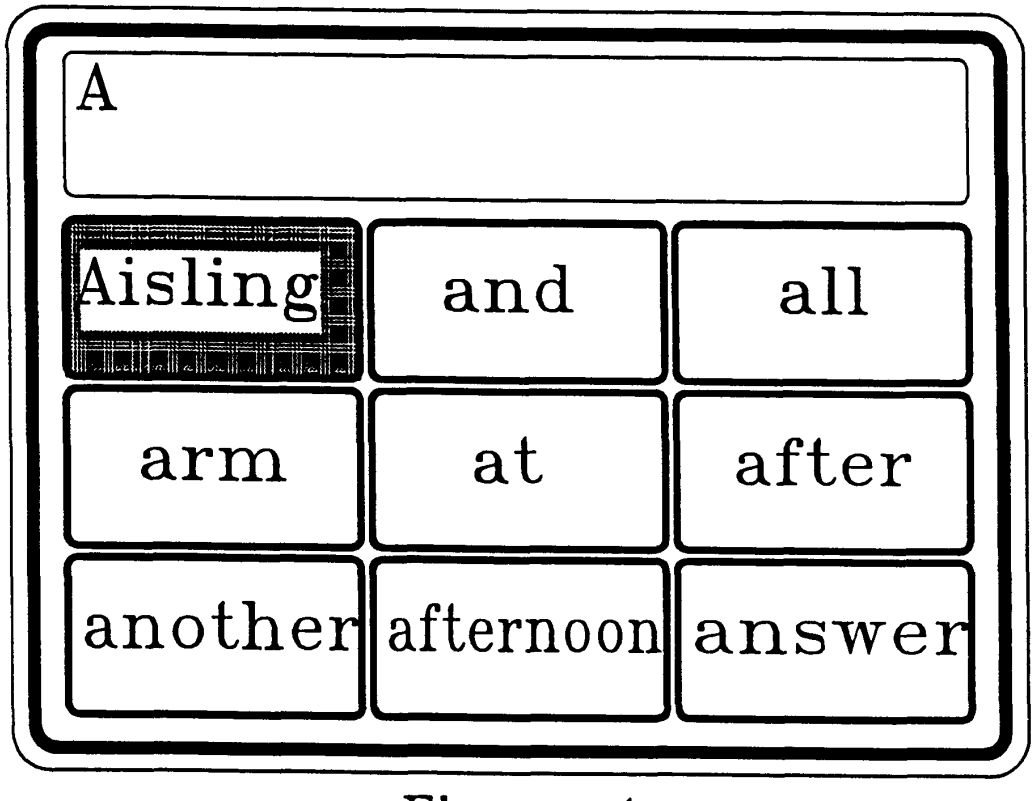

Figure 1.